

Product and Method

This application is a 371 of PCT/GB2003/005102, filed November 21, 2003, the disclosure of which is incorporated herein by reference.

A Sequence Listing on a single CD-ROM was filed with this application (file name: Q87920.ST25.txt). The Sequence Listing contains each of the polynucleotide and polypeptide sequences disclosed herein. The Sequence Listing is incorporated herein by reference.

The present invention relates to oligonucleotide probes, for use in assessing gene transcript levels in a cell, which may be used in analytical techniques, particularly diagnostic techniques. Conveniently the probes are provided in kit form. Different sets of probes may be used in techniques to prepare gene expression patterns and identify, diagnose or monitor different states, such as diseases, conditions or stages thereof. Also provided are methods of identifying suitable probes and their use in methods of the invention.

The identification of quick and easy methods of sample analysis for, for example, diagnostic applications, remains the goal of many researchers. End users seek methods which are cost effective, produce statistically significant results and which may be implemented routinely without the need for highly skilled individuals.

The analysis of gene expression within cells has been used to provide information on the state of those cells and importantly the state of the individual from which the cells are derived. The relative expression of various genes in a cell has been identified as reflecting a particular state within a body. For

example, cancer cells are known to exhibit altered expression of various proteins and the transcripts or the expressed proteins may therefore be used as markers of that disease state.

Thus biopsy tissue may be analysed for the presence of these markers and cells originating from the site of the disease may be identified in other tissues or fluids of the body by the presence of the markers. Furthermore, products of the altered expression may be released into the blood stream and these products may be analysed. In addition cells which have contacted disease cells may be affected by their direct contact with those cells resulting in altered gene expression and their expression or products of expression may be similarly analysed.

However, there are some limitations with these methods. For example, the use of specific tumour markers for identifying cancer suffers from a variety of defects, such as lack of specificity or sensitivity, association of the marker with disease states besides the specific type of cancer, and difficulty of detection in asymptomatic individuals.

In addition to the analysis of one or two marker transcripts or proteins, more recently, gene expression patterns have been analysed. Most of the work involving large-scale gene expression analysis with implications in disease diagnosis has involved clinical samples originating from diseased tissues or cells. For example, several recent publications, which demonstrate that gene expression data can be used to distinguish between similar cancer types, have used clinical samples from diseased tissues or cells (Alon et al. 1999, PNAS, 96, p6745-6750; Golub et al. 1999, Science, 286, p531-537; Alizadeh et al, 2000, Nature, 403, p503-511; Bittner et al., 2000, Nature, 406, p536-540).

However, these methods have relied on analysis of a

sample containing diseased cells or products of those cells or cells which have been contacted by disease cells. Analysis of such samples relies on knowledge of the presence of a disease and its location, which may be difficult in asymptomatic patients. Furthermore, samples can not always be taken from the disease site, e.g. in diseases of the brain.

In a finding of great significance, the present inventors identified the previously untapped potential of all cells within a body to provide information relating to the state of the organism from which the cells were derived. WO98/49342 describes the analysis of the gene expression of cells distant from the site of disease, e.g. peripheral blood collected distant from a cancer site.

This finding is based on the premise that the different parts of an organism's body exist in dynamic interaction with each other. When a disease affects one part of the body, other parts of the body are also affected. The interaction results from a wide spectrum of biochemical signals that are released from the diseased area, affecting other areas in the body. Although, the nature of the biochemical and physiological changes induced by the released signals can vary in the different body parts, the changes can be measured at the level of gene expression and used for diagnostic purposes.

The physiological state of a cell in an organism is determined by the pattern with which genes are expressed in it. The pattern depends upon the internal and external biological stimuli to which said cell is exposed, and any change either in the extent or in the nature of these stimuli can lead to a change in the pattern with which the different genes are expressed in the cell. There is a growing understanding that by analysing the systemic changes in gene expression

patterns in cells in biological samples, it is possible to provide information on the type and nature of the biological stimuli that are acting on them. Thus, for example, by monitoring the expression of a large number of genes in cells in a test sample, it is possible to determine whether their genes are expressed with a pattern characteristic for a particular disease, condition or stage thereof. Measuring changes in gene activities in cells, e.g. from tissue or body fluids is therefore emerging as a powerful tool for disease diagnosis.

Such methods have various advantages. Often, obtaining clinical samples from certain areas in the body that is diseased can be difficult and may involve undesirable invasions in the body, for example biopsy is often used to obtain samples for cancer. In some cases, such as in Alzheimer's disease the diseased brain specimen can only be obtained post-mortem. Furthermore, the tissue specimens which are obtained are often heterogeneous and may contain a mixture of both diseased and non-diseased cells, making the analysis of generated gene expression data both complex and difficult.

It has been suggested that a pool of tumour tissues that appear to be pathogenetically homogeneous with respect to morphological appearances of the tumour may well be highly heterogeneous at the molecular level (Alizadeh, 2000, *supra*), and in fact might contain tumours representing essentially different diseases (Alizadeh, 2000, *supra*; Golub, 1999, *supra*). For the purpose of identifying a disease, condition, or a stage thereof, any method that does not require clinical samples to originate directly from diseased tissues or cells is highly desirable since clinical samples representing a homogeneous mixture of cell types can be obtained from an easily accessible region in the body.

We have now identified a set of probes of

surprising utility for identifying one or more diseases.

Thus, we now describe probes and sets of probes derived from cells which are not disease cells and which have not contacted disease cells, which correspond to genes which exhibit altered expression in normal versus disease individuals, for use in methods of identifying, diagnosing or monitoring certain conditions, particularly diseases or stages thereof.

Thus the invention provides a set of oligonucleotide probes which correspond to genes in a cell whose expression is affected in a pattern characteristic of a particular disease, condition or stage thereof, wherein said genes are systemically affected by said disease, condition or stage thereof. Preferably said genes are metabolic or house-keeping genes and preferably are constitutively moderately or highly expressed. Preferably the genes are moderately or highly expressed in the cells of the sample but not in cells from disease cells or in cells having contacted such disease cells.

Such probes, particularly when isolated from cells distant to the site of disease, do not rely on the development of disease to clinically recognizable levels and allow detection of a disease or condition or stage thereof very early after the onset of said disease or condition, even years before other subjective or objective symptoms appear.

As used herein "systemically" affected genes refers to genes whose expression is affected in the body without direct contact with a disease cell or disease site and the cells under investigation are not disease cells.

"Contact" as referred to herein refers to cells coming into close proximity with one another such that the direct effect of one cell on the other may be observed, e.g. an immune response, wherein these

responses are not mediated by secondary molecules released from the first cell over a large distance to affect the second cell. Preferably contact refers to physical contact, or contact that is as close as is sterically possible, conveniently, cells which contact one another are found in the same unit volume, for example within 1cm^3 .

A "disease cell" is a cell manifesting phenotypic changes and is present at the disease site at some time during its life-span, e.g. a tumour cell at the tumour site or which has disseminated from the tumour, or a brain cell in the case of brain disorders such as Alzheimer's disease.

"Metabolic" or "house-keeping" genes refer to those genes responsible for expressing products involved in cell division and maintenance, e.g. non-immune function related genes.

"Moderately or highly" expressed genes refers to those present in resting cells in a copy number of more than 30-100 copies/cell (assuming an average 3×10^5 mRNA molecules in a cell).

Specific probes having the above described properties are provided herein.

Thus in one aspect, the present invention provides a set of oligonucleotide probes, wherein said set comprises at least 10 oligonucleotides selected from:

an oligonucleotide as described in Table 1 or derived from a sequence described in Table 1, or an oligonucleotide with a complementary sequence, or a functionally equivalent oligonucleotide.

"Table 1" as referred to herein refers to Table 1a and/or Table 1b. Table 1b contains reference to additional clones and sequences as disclosed herein. Similarly Tables 2 and 4 comprise 2 parts, a and b.

The invention also provides one or more oligonucleotide probes, wherein each oligonucleotide

probe is selected from the oligonucleotides listed in Table 1, or derived from a sequence described in Table 1, or a complementary sequence thereof. The use of such probes in products and methods of the invention, form further aspects of the invention. As referred to herein an "oligonucleotide" is a nucleic acid molecule having at least 6 monomers in the polymeric structure, ie. nucleotides or modified forms thereof. The nucleic acid molecule may be DNA, RNA or PNA (peptide nucleic acid) or hybrids thereof or modified versions thereof, e.g. chemically modified forms, e.g. LNA (Locked Nucleic acid), by methylation or made up of modified or non-natural bases during synthesis, providing they retain their ability to bind to complementary sequences. Such oligonucleotides are used in accordance with the invention to probe target sequences and are thus referred to herein also as oligonucleotide probes or simply as probes.

An "oligonucleotide derived from a sequence described in Table 1" (or any other table) refers to a part of a sequence disclosed in that Table (e.g. Table 1-4), which satisfies the requirements of the oligonucleotide probes as described herein, e.g. in length and function. Preferably said parts have the size described hereinafter.

Preferably the oligonucleotide probes forming said set are at least 15 bases in length to allow binding of target molecules. Especially preferably said oligonucleotide probes are from 20 to 200 bases in length, e.g. from 30 to 150 bases, preferably 50-100 bases in length.

As referred to herein the term "complementary sequences" refers to sequences with consecutive complementary bases (ie. T:A, G:C) and which complementary sequences are therefore able to bind to one another through their complementarity.

Reference to "10 oligonucleotides" refers to 10 different oligonucleotides. Whilst a Table 1 oligonucleotide, a Table 1 derived oligonucleotide and their functional equivalent are considered different oligonucleotides, complementary oligonucleotides are not considered different. Preferably however, the at least 10 oligonucleotides are 10 different Table 1 oligonucleotides (or Table 1 derived oligonucleotides or their functional equivalents). Thus said 10 different oligonucleotides are preferably able to bind to 10 different transcripts.

Preferably said oligonucleotides are as described in Table 1 or are derived from a sequence described in Table 1. Especially preferably said oligonucleotides are as described in Table 2 or Table 4 or are derived from a sequence described in either of those tables. Especially preferably the oligonucleotide (or the oligonucleotide derived therefrom) has a high occurrence as defined in Table 3, especially preferably >40%, e.g. >80 or >90, e.g. 100%.

A "set" as described refers to a collection of unique oligonucleotide probes (ie. having a distinct sequence) and preferably consists of less than 1000 oligonucleotide probes, especially less than 500 probes, e.g. preferably from 10 to 500, e.g. 10 to 100, 200 or 300, especially preferably 20 to 100, e.g. 30 to 100 probes. In some cases less than 10 probes may be used, e.g. from 2 to 9 probes, e.g. 5 to 9 probes.

It will be appreciated that increasing the number of probes will prevent the possibility of poor analysis, e.g. misdiagnosis by comparison to other diseases which could similarly alter the expression of the particular genes in question. Other oligonucleotide probes not described herein may also be present, particularly if they aid the ultimate use of the set of oligonucleotide probes. However, preferably said set consists only of

said Table 1 oligonucleotides, Table 1 derived oligonucleotides, complementary sequences or functionally equivalent oligonucleotides, or a sub-set thereof (e.g. of the size as described above), preferably a sub-set for which sequences are provided herein (see Table 1 and its footnote). Especially preferably said set consists only of said Table 1 oligonucleotides, Table 1 derived oligonucleotides, or complementary sequences thereof, or a sub-set thereof.

Multiple copies of each unique oligonucleotide probe, e.g. 10 or more copies, may be present in each set, but constitute only a single probe.

A set of oligonucleotide probes, which may preferably be immobilized on a solid support or have means for such immobilization, comprises the at least 10 oligonucleotide probes selected from those described hereinbefore. Especially preferably said probes are selected from those having high occurrence as described in Table 3 and as mentioned above. As mentioned above, these 10 probes must be unique and have different sequences. Having said this however, two separate probes may be used which recognize the same gene but reflect different splicing events. However oligonucleotide probes which are complementary to, and bind to distinct genes are preferred.

As described herein a "functionally equivalent" oligonucleotide to those described in Table 1 or derived therefrom refers to an oligonucleotide which is capable of identifying the same gene as an oligonucleotide of Table 1 or derived therefrom, ie. it can bind to the same mRNA molecule (or DNA) transcribed from a gene (target nucleic acid molecule) as the Table 1 oligonucleotide or the Table 1 derived oligonucleotide (or its complementary sequence). Preferably said functionally equivalent oligonucleotide is capable of recognizing, ie. binding to the same splicing product as

a Table 1 oligonucleotide or a Table 1 derived oligonucleotide. Preferably said mRNA molecule is the full length mRNA molecule which corresponds to the Table 1 oligonucleotide or the Table 1 derived oligonucleotide.

As referred to herein "capable of binding" or "binding" refers to the ability to hybridize under conditions described hereinafter.

Alternatively expressed, functionally equivalent oligonucleotides (or complementary sequences) have sequence identity or will hybridize, as described hereinafter, to a region of the target molecule to which molecule a Table 1 oligonucleotide or a Table 1 derived oligonucleotide or a complementary oligonucleotide binds. Preferably, functionally equivalent oligonucleotides (or their complementary sequences) hybridize to one of the mRNA sequences which corresponds to a Table 1 oligonucleotide or a Table 1 derived oligonucleotide under the conditions described hereinafter or has sequence identity to a part of one of the mRNA sequences which corresponds to a Table 1 oligonucleotide or a Table 1 derived oligonucleotide. A "part" in this context refers to a stretch of at least 5, e.g. at least 10 or 20 bases, such as from 5 to 100, e.g. 10 to 50 or 15 to 30 bases.

In a particularly preferred aspect, the functionally equivalent oligonucleotide binds to all or a part of the region of a target nucleic acid molecule (mRNA or cDNA) to which the Table 1 oligonucleotide or Table 1 derived oligonucleotide binds. A "target" nucleic acid molecule is the gene transcript or related product e.g. mRNA, or cDNA, or amplified product thereof. Said "region" of said target molecule to which said Table 1 oligonucleotide or Table 1 derived oligonucleotide binds is the stretch over which complementarity exists. At its largest this region is

the whole length of the Table 1 oligonucleotide or Table 1 derived oligonucleotide, but may be shorter if the entire Table 1 sequence or Table 1 derived oligonucleotide is not complementary to a region of the target sequence.

Preferably said part of said region of said target molecule is a stretch of at least 5, e.g. at least 10 or 20 bases, such as from 5 to 100, e.g. 10 to 50 or 15 to 30 bases. This may for example be achieved by said functionally equivalent oligonucleotide having several identical bases to the bases of the Table 1 oligonucleotide or the Table 1 derived oligonucleotide.

These bases may be identical over consecutive stretches, e.g. in a part of the functionally equivalent oligonucleotide, or may be present non-consecutively, but provide sufficient complementarity to allow binding to the target sequence.

Thus in a preferred feature, said functionally equivalent oligonucleotide hybridizes under conditions of high stringency to a Table 1 oligonucleotide or a Table 1 derived oligonucleotide or the complementary sequence thereof. Alternatively expressed, said functionally equivalent oligonucleotide exhibits high sequence identity to all or part of a Table 1 oligonucleotide. Preferably said functionally equivalent oligonucleotide has at least 70% sequence identity, preferably at least 80%, e.g. at least 90, 95, 98 or 99%, to all of a Table 1 oligonucleotide or a part thereof. As used in this context, a "part" refers to a stretch of at least 5, e.g. at least 10 or 20 bases, such as from 5 to 100, e.g. 10 to 50 or 15 to 30 bases, in said Table 1 oligonucleotide. Especially preferably when sequence identity to only a part of said Table 1 oligonucleotide is present, the sequence identity is high, e.g. at least 80% as described above.

Functionally equivalent oligonucleotides which

satisfy the above stated functional requirements include those which are derived from the Table 1 oligonucleotides and also those which have been modified by single or multiple nucleotide base (or equivalent) substitution, addition and/or deletion, but which nonetheless retain functional activity, e.g. bind to the same target molecule as the Table 1 oligonucleotide or the Table 1 derived oligonucleotide from which they are further derived or modified. Preferably said modification is of from 1 to 50, e.g. from 10 to 30, preferably from 1 to 5 bases. Especially preferably only minor modifications are present, e.g. variations in less than 10 bases, e.g. less than 5 base changes.

Within the meaning of "addition" equivalents are included oligonucleotides containing additional sequences which are complementary to the consecutive stretch of bases on the target molecule to which the Table 1 oligonucleotide or the Table 1 derived oligonucleotide binds. Alternatively the addition may comprise a different, unrelated sequence, which may for example confer a further property, e.g. to provide a means for immobilization such as a linker to bind the oligonucleotide probe to a solid support.

Particularly preferred are naturally occurring equivalents such as biological variants, e.g. allelic, geographical or allotypic variants, e.g. oligonucleotides which correspond to a genetic variant, for example as present in a different species.

Functional equivalents include oligonucleotides with modified bases, e.g. using non-naturally occurring bases. Such derivatives may be prepared during synthesis or by post production modification.

"Hybridizing" sequences which bind under conditions of low stringency are those which bind under non-stringent conditions (for example, 6x SSC/50% formamide at room temperature) and remain bound when washed under

conditions of low stringency (2 X SSC, room temperature, more preferably 2 X SSC, 42°C). Hybridizing under high stringency refers to the above conditions in which washing is performed at 2 X SSC, 65°C (where SSC = 0.15M NaCl, 0.015M sodium citrate, pH 7.2).

"Sequence identity" as referred to herein refers to the value obtained when assessed using ClustalW (Thompson et al., 1994, Nucl. Acids Res., 22, p4673-4680) with the following parameters:

Pairwise alignment parameters - Method: accurate, Matrix: IUB, Gap open penalty: 15.00, Gap extension penalty: 6.66;

Multiple alignment parameters - Matrix: IUB, Gap open penalty: 15.00, % identity for delay: 30, Negative matrix: no, Gap extension penalty: 6.66, DNA transitions weighting: 0.5.

Sequence identity at a particular base is intended to include identical bases which have simply been derivatized.

The invention also extends to polypeptides encoded by the mRNA sequence to which a Table 1 oligonucleotide or a Table 1 derived oligonucleotide binds. The invention further extends to antibodies which bind to any of said polypeptides.

As described above, conveniently said set of oligonucleotide probes may be immobilized on one or more solid supports. Single or preferably multiple copies of each unique probe are attached to said solid supports, e.g. 10 or more, e.g. at least 100 copies of each unique probe are present.

One or more unique oligonucleotide probes may be associated with separate solid supports which together form a set of probes immobilized on multiple solid support, e.g. one or more unique probes may be immobilized on multiple beads, membranes, filters, biochips etc. which together form a set of probes, which

together form modules of the kit described hereinafter.

The solid support of the different modules are conveniently physically associated although the signals associated with each probe (generated as described hereinafter) must be separately determinable.

Alternatively, the probes may be immobilized on discrete portions of the same solid support, e.g. each unique oligonucleotide probe, e.g. in multiple copies, may be immobilized to a distinct and discrete portion or region of a single filter or membrane, e.g. to generate an array.

A combination of such techniques may also be used, e.g. several solid supports may be used which each immobilize several unique probes.

The expression "solid support" shall mean any solid material able to bind oligonucleotides by hydrophobic, ionic or covalent bridges.

"Immobilization" as used herein refers to reversible or irreversible association of the probes to said solid support by virtue of such binding. If reversible, the probes remain associated with the solid support for a time sufficient for methods of the invention to be carried out.

Numerous solid supports suitable as immobilizing moieties according to the invention, are well known in the art and widely described in the literature and generally speaking, the solid support may be any of the well-known supports or matrices which are currently widely used or proposed for immobilization, separation etc. in chemical or biochemical procedures. Such materials include, but are not limited to, any synthetic organic polymer such as polystyrene, polyvinylchloride, polyethylene; or nitrocellulose and cellulose acetate; or tosyl activated surfaces; or glass or nylon or any surface carrying a group suited for covalent coupling of nucleic acids. The immobilizing moieties may take the

form of particles, sheets, gels, filters, membranes, microfibre strips, tubes or plates, fibres or capillaries, made for example of a polymeric material e.g. agarose, cellulose, alginate, teflon, latex or polystyrene or magnetic beads. Solid supports allowing the presentation of an array, preferably in a single dimension are preferred, e.g. sheets, filters, membranes, plates or biochips.

Attachment of the nucleic acid molecules to the solid support may be performed directly or indirectly. For example if a filter is used, attachment may be performed by UV-induced crosslinking. Alternatively, attachment may be performed indirectly by the use of an attachment moiety carried on the oligonucleotide probes and/or solid support. Thus for example, a pair of affinity binding partners may be used, such as avidin, streptavidin or biotin, DNA or DNA binding protein (e.g. either the lac I repressor protein or the lac operator sequence to which it binds), antibodies (which may be mono- or polyclonal), antibody fragments or the epitopes or haptens of antibodies. In these cases, one partner of the binding pair is attached to (or is inherently part of) the solid support and the other partner is attached to (or is inherently part of) the nucleic acid molecules.

As used herein an "affinity binding pair" refers to two components which recognize and bind to one another specifically (ie. in preference to binding to other molecules). Such binding pairs when bound together form a complex.

Attachment of appropriate functional groups to the solid support may be performed by methods well known in the art, which include for example, attachment through hydroxyl, carboxyl, aldehyde or amino groups which may be provided by treating the solid support to provide suitable surface coatings. Solid supports presenting

appropriate moieties for attachment of the binding partner may be produced by routine methods known in the art.

Attachment of appropriate functional groups to the oligonucleotide probes of the invention may be performed by ligation or introduced during synthesis or amplification, for example using primers carrying an appropriate moiety, such as biotin or a particular sequence for capture.

Conveniently, the set of probes described hereinbefore is provided in kit form.

Thus viewed from a further aspect the present invention provides a kit comprising a set of oligonucleotide probes as described hereinbefore immobilized on one or more solid supports.

Preferably, said probes are immobilized on a single solid support and each unique probe is attached to a different region of said solid support. However, when attached to multiple solid supports, said multiple solid supports form the modules which make up the kit. Especially preferably said solid support is a sheet, filter, membrane, plate or biochip.

Optionally the kit may also contain information relating to the signals generated by normal or diseased samples (as discussed in more detail hereinafter in relation to the use of the kits), standardizing materials, e.g. mRNA or cDNA from normal and/or diseased samples for comparative purposes, labels for incorporation into cDNA, adapters for introducing nucleic acid sequences for amplification purposes, primers for amplification and/or appropriate enzymes, buffers and solutions. Optionally said kit may also contain a package insert describing how the method of the invention should be performed, optionally providing standard graphs, data or software for interpretation of results obtained when performing the invention.

The use of such kits to prepare a standard diagnostic gene transcript pattern as described hereinafter forms a further aspect of the invention.

The set of probes as described herein have various uses. Principally however they are used to assess the gene expression state of a test cell to provide information relating to the organism from which said cell is derived. Thus the probes are useful in diagnosing, identifying or monitoring a disease or condition or stage thereof in an organism.

Thus in a further aspect the invention provides the use of a set of oligonucleotide probes or a kit as described hereinbefore to determine the gene expression pattern of a cell which pattern reflects the level of gene expression of genes to which said oligonucleotide probes bind, comprising at least the steps of:

- a) isolating mRNA from said cell, which may optionally be reverse transcribed to cDNA;
- b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotide probes or a kit as defined herein; and
- c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce said pattern.

The mRNA and cDNA as referred to in this method, and the methods hereinafter, encompass derivatives or copies of said molecules, e.g. copies of such molecules such as those produced by amplification or the preparation of complementary strands, but which retain the identity of the mRNA sequence, ie. would hybridize to the direct transcript (or its complementary sequence) by virtue of precise complementarity, or sequence identity, over at least a region of said molecule. It will be appreciated that complementarity will not exist over the entire region where techniques have been used which may truncate the transcript or introduce new sequences, e.g. by primer amplification. For

convenience, said mRNA or cDNA is preferably amplified prior to step b). As with the oligonucleotides described herein said molecules may be modified, e.g. by using non-natural bases during synthesis providing complementarity remains. Such molecules may also carry additional moieties such as signalling or immobilizing means.

The various steps involved in the method of preparing such a pattern are described in more detail hereinafter.

As used herein "gene expression" refers to transcription of a particular gene to produce a specific mRNA product (ie. a particular splicing product). The level of gene expression may be determined by assessing the level of transcribed mRNA molecules or cDNA molecules reverse transcribed from the mRNA molecules or products derived from those molecules, e.g. by amplification.

The "pattern" created by this technique refers to information which, for example, may be represented in tabular or graphical form and conveys information about the signal associated with two or more oligonucleotides.

Preferably said pattern is expressed as an array of numbers relating to the expression level associated with each probe.

Preferably, said pattern is established using the following linear model:

$$y = \mathbf{Xb} + \mathbf{f} \quad \text{Equation 1}$$

wherein, \mathbf{X} is the matrix of gene expression data and \mathbf{y} is the response variable, \mathbf{b} is the regression coefficient vector and \mathbf{f} the estimated residual vector. Although many different methods can be used to establish the relationship provided in equation 1, especially preferably the partial Least Squares Regression (PLSR) method is used for establishing the relationship in equation 1.

The probes are thus used to generate a pattern which reflects the gene expression of a cell at the time of its isolation. The pattern of expression is characteristic of the circumstances under which that cells finds itself and depends on the influences to which the cell has been exposed. Thus, a characteristic gene transcript pattern standard or fingerprint (standard probe pattern) for cells from an individual with a particular disease or condition may be prepared and used for comparison to transcript patterns of test cells. This has clear applications in diagnosing, monitoring or identifying whether an organism is suffering from a particular disease, condition or stage thereof.

The standard pattern is prepared by determining the extent of binding of total mRNA (or cDNA or related product), from cells from a sample of one or more organisms with the disease or condition or stage thereof, to the probes. This reflects the level of transcripts which are present which correspond to each unique probe. The amount of nucleic acid material which binds to the different probes is assessed and this information together forms the gene transcript pattern standard of that disease or condition or stage thereof.

Each such standard pattern is characteristic of the disease, condition or stage thereof.

In a further aspect therefore, the present invention provides a method of preparing a standard gene transcript pattern characteristic of a disease or condition or stage thereof in an organism comprising at least the steps of:

- a) isolating mRNA from the cells of a sample of one or more organisms having the disease or condition or stage thereof, which may optionally be reverse transcribed to cDNA;

- b) hybridizing the mRNA or cDNA of step (a) to a

set of oligonucleotides or a kit as described hereinbefore specific for said disease or condition or stage thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and

c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce a characteristic pattern reflecting the level of gene expression of genes to which said oligonucleotides bind, in the sample with the disease, condition or stage thereof.

For convenience, said oligonucleotides are preferably immobilized on one or more solid supports.

The standard pattern for a great number of diseases or conditions and different stages thereof using particular probes may be accumulated in databases and be made available to laboratories on request.

"Disease" samples and organisms as referred to herein refer to organisms (or samples from the same) with an underlying pathological disturbance relative to a normal organism (or sample), in a symptomatic or asymptomatic organism, which may result, for example, from infection or an acquired or congenital genetic imperfection. Such organisms are known to have, or which exhibit, the disease or condition or stage thereof under study.

A "condition" refers to a state of the mind or body of an organism which has not occurred through disease, e.g. the presence of an agent in the body such as a toxin, drug or pollutant, or pregnancy.

"Stages" thereof refer to different stages of the disease or condition which may or may not exhibit particular physiological or metabolic changes, but do exhibit changes at the genetic level which may be detected as altered gene expression. It will be appreciated that during the course of a disease or condition the expression of different transcripts may

vary. Thus at different stages, altered expression may not be exhibited for particular transcripts compared to "normal" samples. However, combining information from several transcripts which exhibit altered expression at one or more stages through the course of the disease or condition can be used to provide a characteristic pattern which is indicative of a particular stage of the disease or condition. Thus for example different stages in cancer, e.g. pre-stage I, stage I, stage II, II or IV can be identified.

"Normal" as used herein refers to organisms or samples which are used for comparative purposes. Preferably, these are "normal" in the sense that they do not exhibit any indication of, or are not believed to have, any disease or condition that would affect gene expression, particularly in respect of the disease for which they are to be used as the normal standard. However, it will be appreciated that different stages of a disease or condition may be compared and in such cases, the "normal" sample may correspond to the earlier stage of the disease or condition.

As used herein a "sample" refers to any material obtained from the organism, e.g. human or non-human animal under investigation which contains cells and includes, tissues, body fluid or body waste or in the case of prokaryotic organisms, the organism itself. "Body fluids" include blood, saliva, spinal fluid, semen, lymph. "Body waste" includes urine, expectorated matter (pulmonary patients), faeces etc. "Tissue samples" include tissue obtained by biopsy, by surgical interventions or by other means e.g. placenta. Preferably however, the samples which are examined are from areas of the body not apparently affected by the disease or condition. The cells in such samples are not disease cells, e.g. cancer cells, have not been in contact with such disease cells and do not originate

from the site of the disease or condition. The "site of disease" is considered to be that area of the body which manifests the disease in a way which may be objectively determined, e.g. a tumour or area of inflammation. Thus for example peripheral blood may be used for the diagnosis of non-haematopoietic cancers, and the blood does not require the presence of malignant or disseminated cells from the cancer in the blood. Similarly in diseases of the brain, in which no diseased cells are found in the blood due to the blood:brain barrier, peripheral blood may still be used in the methods of the invention.

It will however be appreciated that the method of preparing the standard transcription pattern and other methods of the invention are also applicable for use on living parts of eukaryotic organisms such as cell lines and organ cultures and explants. As used herein, reference to "corresponding" sample etc. refers to cells preferably from the same tissue, body fluid or body waste, but also includes cells from tissue, body fluid or body waste which are sufficiently similar for the purposes of preparing the standard or test pattern. When used in reference to genes "corresponding" to the probes, this refers to genes which are related by sequence (which may be complementary) to the probes although the probes may reflect different splicing products of expression.

"Assessing" as used herein refers to both quantitative and qualitative assessment which may be determined in absolute or relative terms.

The invention may be put into practice as follows.

To prepare a standard transcript pattern for a particular disease, condition or stage thereof, sample mRNA is extracted from the cells of tissues, body fluid or body waste according to known techniques (see for

example Sambrook et. al. (1989), Molecular Cloning : A laboratory manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.) from a diseased individual or organism.

Owing to the difficulties in working with RNA, the RNA is preferably reverse transcribed at this stage to form first strand cDNA. Cloning of the cDNA or selection from, or using, a cDNA library is not however necessary in this or other methods of the invention. Preferably, the complementary strands of the first strand cDNAs are synthesized, ie. second strand cDNAs, but this will depend on which relative strands are present in the oligonucleotide probes. The RNA may however alternatively be used directly without reverse transcription and may be labelled if so required.

Preferably the cDNA strands are amplified by known amplification techniques such as the polymerase chain reaction (PCR) by the use of appropriate primers. Alternatively, the cDNA strands may be cloned with a vector, used to transform a bacteria such as E. coli which may then be grown to multiply the nucleic acid molecules. When the sequence of the cDNAs are not known, primers may be directed to regions of the nucleic acid molecules which have been introduced. Thus for example, adapters may be ligated to the cDNA molecules and primers directed to these portions for amplification of the cDNA molecules. Alternatively, in the case of eukaryotic samples, advantage may be taken of the polyA tail and cap of the RNA to prepare appropriate primers.

To produce the standard diagnostic gene transcript pattern or fingerprint for a particular disease or condition or stage thereof, the above described oligonucleotide probes are used to probe mRNA or cDNA of the diseased sample to produce a signal for hybridization to each particular oligonucleotide probe species, ie. each unique probe. A standard control gene

transcript pattern may also be prepared if desired using mRNA or cDNA from a normal sample. Thus, mRNA or cDNA is brought into contact with the oligonucleotide probe under appropriate conditions to allow hybridization.

When multiple samples are probed, this may be performed consecutively using the same probes, e.g. on one or more solid supports, ie. on probe kit modules, or by simultaneously hybridizing to corresponding probes, e.g. the modules of a corresponding probe kit.

To identify when hybridization occurs and obtain an indication of the number of transcripts/cDNA molecules which become bound to the oligonucleotide probes, it is necessary to identify a signal produced when the transcripts (or related molecules) hybridize (e.g. by detection of double stranded nucleic acid molecules or detection of the number of molecules which become bound, after removing unbound molecules, e.g. by washing).

In order to achieve a signal, either or both components which hybridize (ie. the probe and the transcript) carry or form a signalling means or a part thereof. This "signalling means" is any moiety capable of direct or indirect detection by the generation or presence of a signal. The signal may be any detectable physical characteristic such as conferred by radiation emission, scattering or absorption properties, magnetic properties, or other physical properties such as charge, size or binding properties of existing molecules (e.g. labels) or molecules which may be generated (e.g. gas emission etc.). Techniques are preferred which allow signal amplification, e.g. which produce multiple signal events from a single active binding site, e.g. by the catalytic action of enzymes to produce multiple detectable products.

Conveniently the signalling means may be a label which itself provides a detectable signal. Conveniently this may be achieved by the use of a radioactive or

other label which may be incorporated during cDNA production, the preparation of complementary cDNA strands, during amplification of the target mRNA/cDNA or added directly to target nucleic acid molecules.

Appropriate labels are those which directly or indirectly allow detection or measurement of the presence of the transcripts/cDNA. Such labels include for example radiolabels, chemical labels, for example chromophores or fluorophores (e.g. dyes such as fluorescein and rhodamine), or reagents of high electron density such as ferritin, haemocyanin or colloidal gold.

Alternatively, the label may be an enzyme, for example peroxidase or alkaline phosphatase, wherein the presence of the enzyme is visualized by its interaction with a suitable entity, for example a substrate. The label may also form part of a signalling pair wherein the other member of the pair is found on, or in close proximity to, the oligonucleotide probe to which the transcript/cDNA binds, for example, a fluorescent compound and a quench fluorescent substrate may be used.

A label may also be provided on a different entity, such as an antibody, which recognizes a peptide moiety attached to the transcripts/cDNA, for example attached to a base used during synthesis or amplification.

A signal may be achieved by the introduction of a label before, during or after the hybridization step. Alternatively, the presence of hybridizing transcripts may be identified by other physical properties, such as their absorbance, and in which case the signalling means is the complex itself.

The amount of signal associated with each oligonucleotide probe is then assessed. The assessment may be quantitative or qualitative and may be based on binding of a single transcript species (or related cDNA or other products) to each probe, or binding of multiple transcript species to multiple copies of each unique

probe. It will be appreciated that quantitative results will provide further information for the transcript fingerprint of the disease which is compiled. This data may be expressed as absolute values (in the case of macroarrays) or may be determined relative to a particular standard or reference e.g. a normal control sample.

Furthermore it will be appreciated that the standard diagnostic gene pattern transcript may be prepared using one or more disease samples (and normal samples if used) to perform the hybridization step to obtain patterns not biased towards a particular individual's variations in gene expression.

The use of the probes to prepare standard patterns and the standard diagnostic gene transcript patterns thus produced for the purpose of identification or diagnosis or monitoring of a particular disease or condition or stage thereof in a particular organism forms a further aspect of the invention.

Once a standard diagnostic fingerprint or pattern has been determined for a particular disease or condition using the selected oligonucleotide probes, this information can be used to identify the presence, absence or extent or stage of that disease or condition in a different test organism or individual.

To examine the gene expression pattern of a test sample, a test sample of tissue, body fluid or body waste containing cells, corresponding to the sample used for the preparation of the standard pattern, is obtained from a patient or the organism to be studied. A test gene transcript pattern is then prepared as described hereinbefore as for the standard pattern.

In a further aspect therefore, the present invention provides a method of preparing a test gene transcript pattern comprising at least the steps of:

- a) isolating mRNA from the cells of a sample of

said test organism, which may optionally be reverse transcribed to cDNA;

b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotides or a kit as described hereinbefore specific for a disease or condition or stage thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and

c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce said pattern reflecting the level of gene expression of genes to which said oligonucleotides bind, in said test sample.

This test pattern may then be compared to one or more standard patterns to assess whether the sample contains cells having the disease, condition or stage thereof.

Thus viewed from a further aspect the present invention provides a method of diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism, comprising the steps of:

- a) isolating mRNA from the cells of a sample of said organism, which may optionally be reverse transcribed to cDNA;
- b) hybridizing the mRNA or cDNA of step (a) to a set of oligonucleotides or a kit as described hereinbefore specific for said disease or condition or stage thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation;
- c) assessing the amount of mRNA or cDNA hybridizing to each of said probes to produce a characteristic pattern reflecting the level of gene expression of genes to which said oligonucleotides bind, in said sample; and
- d) comparing said pattern to a standard diagnostic pattern prepared according to the

method of the invention using a sample from an organism corresponding to the organism and sample under investigation to determine the presence of said disease or condition or a stage thereof in the organism under investigation.

The method up to and including step c) is the preparation of a test pattern as described above.

As referred to herein, "diagnosis" refers to determination of the presence or existence of a disease or condition or stage thereof in an organism.

"Monitoring" refers to establishing the extent of a disease or condition, particularly when an individual is known to be suffering from a disease or condition, for example to monitor the effects of treatment or the development of a disease or condition, e.g. to determine the suitability of a treatment or provide a prognosis.

The presence of the disease or condition or stage thereof may be determined by determining the degree of correlation between the standard and test samples' patterns. This necessarily takes into account the range of values which are obtained for normal and diseased samples. Although this can be established by obtaining standard deviations for several representative samples binding to the probes to develop the standard, it will be appreciated that single samples may be sufficient to generate the standard pattern to identify a disease if the test sample exhibits close enough correlation to that standard. Conveniently, the presence, absence, or extent of a disease or condition or stage thereof in a test sample can be predicted by inserting the data relating to the expression level of informative probes in test sample into the standard diagnostic probe pattern established according to equation 1.

Data generated using the above mentioned methods may be analysed using various techniques from the most

basic visual representation (e.g. relating to intensity) to more complex data manipulation to identify underlying patterns which reflect the interrelationship of the level of expression of each gene to which the various probes bind, which may be quantified and expressed mathematically. Conveniently, the raw data thus generated may be manipulated by the data processing and statistical methods described hereinafter, particularly normalizing and standardizing the data and fitting the data to a classification model to determine whether said test data reflects the pattern of a particular disease, condition or stage thereof.

The methods described herein may be used to identify, monitor or diagnose a disease, condition or ailment or its stage or progression, for which the oligonucleotide probes are informative. "Informative" probes as described herein, are those which reflect genes which have altered expression in the diseases or conditions in question, or particular stages thereof. Probes of the invention may not be sufficiently informative for diagnostic purposes when used alone, but are informative when used as one of several probes to provide a characteristic pattern, e.g. in a set as described hereinbefore.

Preferably said probes correspond to genes which are systemically affected by said disease, condition or stage thereof. Especially preferably said genes, from which transcripts are derived which bind to probes of the invention, are metabolic or house-keeping genes and preferably are moderately or highly expressed. The advantage of using probes directed to moderately or highly expressed genes is that smaller clinical samples are required for generating the necessary gene expression data set, e.g. less than 1ml blood samples.

Furthermore, it has been found that such genes which are already being actively transcribed tend to be

more prone to being influenced, in a positive or negative way, by new stimuli. In addition, since transcripts are already being produced at levels which are generally detectable, small changes in those levels are readily detectable as for example, a certain detectable threshold does not need to be reached.

In preferred methods of the invention, the set of probes of the invention are informative for a variety of different diseases, conditions or stages thereof. A sub-set of the probes disclosed herein may be used for diagnosis, identification or monitoring a particular disease, condition or stage thereof. Thus the probes may be used to diagnose or identify or monitor any condition, ailment, disease or reaction that leads to the relative increase or decrease in the activity of informative genes of any or all eukaryotic or prokaryotic organisms regardless of whether these changes have been caused by the influence of bacteria, virus, prions, parasites, fungi, radiation, natural or artificial toxins, drugs or allergens, including mental conditions due to stress, neurosis, psychosis or deteriorations due to the ageing of the organism, and conditions or diseases of unknown cause, providing a sub-set of the probes as described herein are informative for said disease or condition or stage thereof.

Such diseases include those which result in metabolic or physiological changes, such as fever-associated diseases such as influenza or malaria. Other diseases which may be detected include for example yellow fever, sexually transmitted diseases such as gonorrhea, fibromyalgia, candida-related complex, cancer (for example of the stomach, lung, breast, prostate gland, bowel, skin, colon, ovary etc), Alzheimer's disease, disease caused by retroviruses such as HIV, senile dementia, multiple sclerosis and Creutzfeldt-

Jakob disease to mention a few.

The invention may also be used to identify patients with psychiatric or psychosomatic diseases such as schizophrenia and eating disorders. Of particular importance is the use of this method to detect diseases, conditions, or stages thereof, which are not readily detectable by known diagnostic methods, such as HIV which is generally not detectable using known techniques 1 to 4 months following infection. Conditions which may be identified include for example drug abuse, such as the use of narcotics, alcohol, steroids or performance enhancing drugs.

Preferably said disease to be identified or monitored is a cancer or a degenerative brain disorder (such as Alzheimer's or Parkinson's disease).

In particular, a set of oligonucleotide probes, wherein said set comprises at least 10 oligonucleotides selected from:

an oligonucleotide as described in Table 4 or an oligonucleotide derived therefrom or an oligonucleotide with a complementary sequence, or a functionally equivalent oligonucleotide, may be used for diagnosis or identification or monitoring the progression of Alzheimer's disease. Similarly Table 2 probes and Table 2 derived probes and their functional equivalents may be used to diagnose, identify or monitor the progression of breast cancer. Especially preferably the probes used for breast cancer analysis are selected based on their occurrence as set forth in Table 3 and as described hereinbefore.

The diagnostic method may be used alone as an alternative to other diagnostic techniques or in addition to such techniques. For example, methods of the invention may be used as an alternative or additive diagnostic measure to diagnosis using imaging techniques such as Magnetic Resonance Image (MRI), ultrasound

imaging, nuclear imaging or X-ray imaging, for example in the identification and/or diagnosis of tumours.

The methods of the invention may be performed on cells from prokaryotic or eukaryotic organisms which may be any eukaryotic organisms such as human beings, other mammals and animals, birds, insects, fish and plants, and any prokaryotic organism such as a bacteria.

Preferred non-human animals on which the methods of the invention may be conducted include, but are not limited to mammals, particularly primates, domestic animals, livestock and laboratory animals. Thus preferred animals for diagnosis include mice, rats, guinea pigs, cats, dogs, pigs, cows, goats, sheep, horses. Particularly preferably the disease state or condition of humans is diagnosed, identified or monitored.

As described above, the sample under study may be any convenient sample which may be obtained from an organism. Preferably however, as mentioned above, the sample is obtained from a site distant to the site of disease and the cells in such samples are not disease cells, have not been in contact with such cells and do not originate from the site of the disease or condition.

In such cases, although preferably absent, the sample may contain cells which do not fulfil these criteria. However, since the probes of the invention are concerned with transcripts whose expression is altered in cells which do satisfy these criteria, the probes are specifically directed to detecting changes in transcript levels in those cells even if in the presence of other, background cells.

It has been found that the cells from such samples show significant and informative variations in the gene expression of a large number of genes. Thus, the same probe (or several probes) may be found to be informative in determinations regarding two or more diseases,

conditions or stages thereof by virtue of the particular level of transcripts binding to that probe or the interrelationship of the extent of binding to that probe relative to other probes. As a consequence, it is possible to use a relatively small number of probes for screening for multiple disorders or diseases. This has consequences with regard to the selection of probes, discussed in relation to random identification of probes hereinafter, but also for the use of a single set of probes for more than one diagnosis. Table 9 which represents preferred probes of the invention discloses probes which are informative for both Alzheimer's and breast cancer.

Thus, the present invention also provides sets of probes for diagnosing, identifying or monitoring two or more diseases, conditions or stages thereof, wherein at least one of said probes is suitable for said diagnosing, identifying or monitoring at least two of said diseases, conditions or stages thereof, and kits and methods of using the same. Preferably at least 5 probes, e.g. from 5 to 15 probes, are used in at least two diagnoses.

Thus, in a further preferred aspect, the present invention provides a method of diagnosis or identification or monitoring as described hereinbefore for the diagnosis, identification or monitoring of two or more diseases, conditions or stages thereof in an organism, wherein said test pattern produced in step c) of the diagnostic method is compared in step d) to at least two standard diagnostic patterns prepared as described previously, wherein each standard diagnostic pattern is a pattern generated for a different disease or condition or stage thereof.

Whilst in a preferred aspect the methods of assessment concern the development of a gene transcript pattern from a test sample and comparison of the same to

a standard pattern, the elevation or depression of expression of certain markers may also be examined by examining the products of expression and the level of those products. Thus a standard pattern in relation to the expressed product may be generated.

In such methods the levels of expression of a set of polypeptides encoded by the gene to which an oligonucleotide of Table 1 or a Table 1 derived oligonucleotide, binds, are analysed.

Various diagnostic methods may be used to assess the amount of polypeptides (or fragments thereof) which are present. The presence or concentration of polypeptides may be examined, for example by the use of a binding partner to said polypeptide (e.g. an antibody), which may be immobilized, to separate said polypeptide from the sample and the amount of polypeptide may then be determined.

"Fragments" of the polypeptides refers to a domain or region of said polypeptide, e.g. an antigenic fragment, which is recognizable as being derived from said polypeptide to allow binding of a specific binding partner. Preferably such a fragment comprises a significant portion of said polypeptide and corresponds to a product of normal post-synthesis processing. Thus in a further aspect the present invention provides a method of preparing a standard gene transcript pattern characteristic of a disease or condition or stage thereof in an organism comprising at least the steps of:

a) releasing target polypeptides from a sample of one or more organisms having the disease or condition or stage thereof;

b) contacting said target polypeptides with one or more binding partners, wherein each binding partner is specific to a marker polypeptide (or a fragment thereof) encoded by the gene to which an oligonucleotide of Table 1 (or derived from a sequence described in Table 1)

binds, to allow binding of said binding partners to said target polypeptides, wherein said marker polypeptides are specific for said disease or condition thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and

c) assessing the target polypeptide binding to said binding partners to produce a characteristic pattern reflecting the level of gene expression of genes which express said marker polypeptides, in the sample with the disease, condition or stage thereof.

As used herein "target polypeptides" refer to those polypeptides present in a sample which are to be detected and "marker polypeptides" are polypeptides which are encoded by the genes to which Table 1 oligonucleotides or Table 1 derived oligonucleotides bind. The target and marker polypeptides are identical or at least have areas of high similarity, e.g. epitopic regions to allow recognition and binding of the binding partner.

"Release" of the target polypeptides refers to appropriate treatment of a sample to provide the polypeptides in a form accessible for binding of the binding partners, e.g. by lysis of cells where these are present. The samples used in this case need not necessarily comprise cells as the target polypeptides may be released from cells into the surrounding tissue or fluid, and this tissue or fluid may be analysed, e.g. urine or blood. Preferably however the preferred samples as described herein are used. "Binding partners" comprise the separate entities which together make an affinity binding pair as described above, wherein one partner of the binding pair is the target or marker polypeptide and the other partner binds specifically to that polypeptide, e.g. an antibody.

Various arrangements may be envisaged for detecting the amount of binding pairs which form. In its simplest

form, a sandwich type assay e.g. an immunoassay such as an ELISA, may be used in which an antibody specific to the polypeptide and carrying a label (as described elsewhere herein) may be bound to the binding pair (e.g. the first antibody:polypeptide pair) and the amount of label detected.

Other methods as described herein may be similarly modified for analysis of the protein product of expression rather than the gene transcript and related nucleic acid molecules.

Thus a further aspect of the invention provides a method of preparing a test gene transcript pattern comprising at least the steps of:

- a) releasing target polypeptides from a sample of said test organism;

- b) contacting said target polypeptides with one or more binding partners, wherein each binding partner is specific to a marker polypeptide (or a fragment thereof) encoded by the gene to which an oligonucleotide of Table 1 (or derived from a sequence described in Table 1) binds, to allow binding of said binding partners to said target polypeptides, wherein said marker polypeptides are specific for said disease or condition thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and

- c) assessing the target polypeptide binding to said binding partners to produce a characteristic pattern reflecting the level of gene expression of genes which express said marker polypeptides, in said test sample.

A yet further aspect of the invention provides a method of diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism comprising the steps of:

- a) releasing target polypeptides from a sample of said organism;

- b) contacting said target polypeptides with one or

more binding partners, wherein each binding partner is specific to a marker polypeptide (or a fragment thereof) encoded by the gene to which an oligonucleotide of Table 1 (or derived from a sequence described in Table 1) binds, to allow binding of said binding partners to said target polypeptides, wherein said marker polypeptides are specific for said disease or condition thereof in an organism and sample thereof corresponding to the organism and sample thereof under investigation; and

c) assessing the target polypeptide binding to said binding partners to produce a characteristic pattern reflecting the level of gene expression of genes which express said marker polypeptides in said sample; and

d) comparing said pattern to a standard diagnostic pattern prepared as described hereinbefore using a sample from an organism corresponding to the organism and sample under investigation to determine the degree of correlation indicative of the presence of said disease or condition or a stage thereof in the organism under investigation.

The methods of generating standard and test patterns and diagnostic techniques rely on the use of informative oligonucleotide probes to generate the gene expression data. In some cases it will be necessary to select these informative probes for a particular method, e.g. to diagnose a particular disease, from a selection of available probes, e.g. the probes described hereinbefore (the Table 1 oligonucleotides, the Table 1 derived oligonucleotides, their complementary sequences and functionally equivalent oligonucleotides). The following methodology describes a convenient method for identifying such informative probes, or more particularly how to select a suitable sub-set of probes from the probes described herein.

Probes for the analysis of a particular disease or condition or stage thereof, may be identified in a

number of ways known in the prior art, including by differential expression or by library subtraction (see for example WO98/49342). As described hereinafter, in view of the high information content of most transcripts, as a starting point one may also simply analyse a random sub-set of mRNA or cDNA species and pick the most informative probes from that sub-set. The following method describes the use of immobilized oligonucleotide probes (e.g. the probes of the invention) to which mRNA (or related molecules) from different samples is bound to identify which probes are the most informative to identify a particular type of sample, e.g. a disease sample.

The immobilized probes can be derived from various unrelated or related organisms; the only requirement is that the immobilized probes should bind specifically to their homologous counterparts in test organisms. Probes can also be derived from commercially available or public databases and immobilized on solid supports or, as mentioned above, they can be randomly picked and isolated from a cDNA library and immobilized on a solid support.

The length of the probes immobilised on the solid support should be long enough to allow for specific binding to the target sequences. The immobilised probes can be in the form of DNA, RNA or their modified products or PNAs (peptide nucleic acids). Preferably, the probes immobilised should bind specifically to their homologous counterparts representing highly and moderately expressed genes in test organisms. Conveniently the probes which are used are the probes described herein.

The gene expression pattern of cells in biological samples can be generated using prior art techniques such as microarray or macroarray as described below or using methods described herein. Several technologies have now

been developed for monitoring the expression level of a large number of genes simultaneously in biological samples, such as, high-density oligoarrays (Lockhart et al., 1996, Nat. Biotech., 14, p1675-1680), cDNA microarrays (Schena et al, 1995, Science, 270, p467-470) and cDNA macroarrays (Maier E et al., 1994, Nucl. Acids Res., 22, p3423-3424; Bernard et al., 1996, Nucl. Acids Res., 24, p1435-1442).

In high-density oligoarrays and cDNA microarrays, hundreds and thousands of probe oligonucleotides or cDNAs, are spotted onto glass slides or nylon membranes, or synthesized on biochips. The mRNA isolated from the test and reference samples are labelled by reverse transcription with a red or green fluorescent dye, mixed, and hybridised to the microarray. After washing, the bound fluorescent dyes are detected by a laser, producing two images, one for each dye. The resulting ratio of the red and green spots on the two images provides the information about the changes in expression levels of genes in the test and reference samples. Alternatively, single channel or multiple channel microarray studies can also be performed.

In cDNA macroarray, different cDNAs are spotted on a solid support such as nylon membranes in excess in relation to the amount of test mRNA that can hybridise to each spot. mRNA isolated from test samples is radio-labelled by reverse transcription and hybridised to the immobilised probe cDNA. After washing, the signals associated with labels hybridising specifically to immobilised probe cDNA are detected and quantified. The data obtained in macroarray contains information about the relative levels of transcripts present in the test samples. Whilst macroarrays are only suitable to monitor the expression of a limited number of genes, microarrays can be used to monitor the expression of several thousand genes simultaneously and is, therefore,

a preferred choice for large-scale gene expression studies.

A macroarray technique for generating the gene expression data set has been used to illustrate the probe identification method described herein. For this purpose, mRNA is isolated from samples of interest and used to prepare labelled target molecules, e.g. mRNA or cDNA as described above. The labelled target molecules are then hybridised to probes immobilised on the solid support. Various solid supports can be used for the purpose, as described previously. Following hybridization, unbound target molecules are removed and signals from target molecules hybridizing to immobilised probes quantified. If radio labelling is performed, PhosphoImager can be used to generate an image file that can be used to generate a raw data set. Depending on the nature of label chosen for labelling the target molecules, other instruments can also be used, for example, when fluorescence is used for labelling, a FluoroImager can be used to generate an image file from the hybridised target molecules.

The raw data corresponding to mean intensity, median intensity, or volume of the signals in each spot can be acquired from the image file using commercially available software for image analysis. However, the acquired data needs to be corrected for background signals and normalized prior to analysis, since, several factors can affect the quality and quantity of the hybridising signals. For example, variations in the quality and quantity of mRNA isolated from sample to sample, subtle variations in the efficiency of labelling target molecules during each reaction, and variations in the amount of unspecific binding between different macroarrays can all contribute to noise in the acquired data set that must be corrected for prior to analysis.

Background correction can be performed in several

ways. The lowest pixel intensity within a spot can be used for background subtraction or the mean or median of the line of pixels around the spots' outline can be used for the purpose. One can also define an area representing the background intensity based on the signals generated from negative controls and use the average intensity of this area for background subtraction.

The background corrected data can then be transformed for stabilizing the variance in the data structure and normalized for the differences in probe intensity. Several transformation techniques have been described in the literature and a brief overview can be found in Cui, Kerr and Churchill (<http://www.jax.org/research/churchill/research/expression/Cui-Transform.pdf>). Normalization can be performed by dividing the intensity of each spot with the collective intensity, average intensity or median intensity of all the spots in a macroarray or a group of spots in a macroarray in order to obtain the relative intensity of signals hybridising to immobilised probes in a macroarray. Several methods have been described for normalizing gene expression data (Richmond and Somerville, 2000, *Current Opin. Plant Biol.*, 3, p108-116; Finkelstein et al., 2001, In "Methods of Microarray Data Analysis. Papers from CAMDA, Eds. Lin & Johnson, Kluwer Academic, p57-68; Yang et al., 2001, In "Optical Technologies and Informatics", Eds. Bittner, Chen, Dorsel & Dougherty, *Proceedings of SPIE*, 4266, p141-152; Dudoit et al, 2000, *J. Am. Stat. Ass.*, 97, p77-87; Alter et al 2000, *supra*; Newton et al., 2001, *J. Comp. Biol.*, 8, p37-52). Generally, a scaling factor or function is first calculated to correct the intensity effect and then used for normalising the intensities. The use of external controls has also been suggested for improved normalization.

One other major challenge encountered in large-scale gene expression analysis is that of standardization of data collected from experiments performed at different times. We have observed that gene expression data for samples acquired in the same experiment can be efficiently compared following background correction and normalization. However, the data from samples acquired in experiments performed at different times requires further standardization prior to analysis. This is because subtle differences in experimental parameters between different experiments, for example, differences in the quality and quantity of mRNA extracted at different times, differences in time used for target molecule labelling, hybridization time or exposure time, can affect the measured values. Also, factors such as the nature of the sequence of transcripts under investigation (their GC content) and their amount in relation to the each other determines how they are affected by subtle variations in the experimental processes. They determine, for example, how efficiently first strand cDNAs, corresponding to a particular transcript, are transcribed and labelled during first strand synthesis, or how efficiently the corresponding labelled target molecules bind to their complementary sequences during hybridization. Batch to batch difference in the printing process is also a major factor for variation in the generated expression data.

Failure to properly address and rectify for these influences leads to situations where the differences between the experimental series may overshadow the main information of interest contained in the gene expression data set, i.e. the differences within the combined data from the different experimental series. Figure 1 provides one such example showing a classification based on Principal Component Analysis (PCA) of combined data from two experimental series where the main goal is to

distinguish between Alzheimer/non-Alzheimer patients.

PCA (also known as singular value decomposition) is a technique for studying interdependencies and underlying relationships of a set of variables. The data are modelled in terms of a few significant factors or principal components (PC's), plus residuals. The PC's contain the main phenomena and define the systematic variability present in the data, while the residuals represent the variability interpreted as noise. Details on PCA can be found in Jolliffe (1986, *Principal Component Analysis*, Springer-Verlag, NY), and Jackson (1991, *A User's Guide to Principal Components*, Wiley, NY). The results of Figure 1 show that two clusters are formed representing the data from two experimental series rather than the Alzheimer/non-Alzheimer differentiation. There were eight samples in common between the two series of experiments, which ideally should have fallen on top of, or in near proximity to, each other if appropriately standardized.

We have now found that gene expression data between different experiments can be efficiently standardized by including a subset of samples from one experimental series in the next experimental series and using a direct standardization method (DS), originally described by Wang and Kowalski (*Anal. Chem.*, 1991, 63, p2750 and *J. Chemometrics*, 1991, 5, p129-145). Although the method of DS is well known in the field of analytical chemistry, it remains undescribed and unused in the field of gene expression data analysis.

In DS, the secondary data representing for example experimental series 2 (secondary measurements, R_2) are corrected to match the data measured on the primary measurements representing data from series 1 (R_1), while the calibration model remains unchanged. In DS, response matrices for both experimental series are

related to each other by a transformation matrix F , i.e.

$$R_1 = R_2 F \quad (1)$$

Where F is a square matrix dimensioned gene by gene. From (1), the transformation matrix is calculated as:

$$F = R_2^+ R_1 \quad (2)$$

The transformation matrix F in equation (2) is calculated using a relatively small subset of samples which are measured on both the master primary and the secondary series of data.

Finally, the response of the unknown sample measured on the secondary series $r_{2,un}^T$, is standardized

to the response vector $r_{1,un}^T$ expected from the primary series

$$\hat{r}_{1,un}^T = r_{21,un}^T \hat{F} \quad (3)$$

From the preceding equation it can be seen that the column i of the transformation matrix contains the multiplication factors for a set of genes measured in the secondary series to obtain the intensity at spot i of the corrected series.

The number of samples that are repeated in the experimental series, R_1 and R_2 , should be equal to their ranks, which in this case is equal to the number of principal components retained for explaining the variation in the R_1 and R_2 . For example, if three principal components are retained for explaining the variation in the data set, a minimum of three samples should be repeated between R_1 and R_2 . The samples that should be repeated between different series should ideally be those that exhibit high leverages in the gene

expression pattern. At times, two samples may suffice, while at other times, more than two samples should be ideally be included for good representativity. In some cases, the samples selected can be the same in all the experimental series to be compared (reference samples), while in other cases, representative samples can be selected sequentially by analyzing the expression pattern after each experiment. The selected samples with high leverages are then included in the next experimental series. The results of using Direct Standardization are shown in Figure 1.

Another approach for normalizing and standardizing the gene expression data set is to hybridize each DNA array with target molecules prepared from a test sample and an equal amount of labelled target molecules prepared from representative reference samples. In order to measure the intensity of labelled target molecules hybridizing to the immobilized probes it is necessary that the labelled molecules are prepared from test and reference samples using different labels, for example, different fluorescent dyes can be used for preparing the labelled material. The labelled molecules prepared from reference samples can be added to the hybridization solution together with the labelled material prepared from test samples. A data file from each array representing the expression pattern of different genes in the test sample and reference samples can then be obtained, normalized and standardized by the direct standardization method as described above. An instant advantage of including the differentially labelled target molecules from reference samples during hybridization is that it enables an efficient comparison of new test samples to the data sets already stored in a database.

Monitoring the expression of a large number of genes in several samples leads to the generation of a

large amount of data that is too complex to be easily interpreted. Several unsupervised and supervised multivariate data analysis techniques have already been shown to be useful in extracting meaningful biological information from these large data sets. Cluster analysis is by far the most commonly used technique for gene expression analysis, and has been performed to identify genes that are regulated in a similar manner, and or identifying new/unknown tumour classes using gene expression profiles (Eisen et al., 1998, PNAS, 95, p14863-14868, Alizadeh et al. 2000, supra, Perou et al. 2000, Nature, 406, p747-752; Ross et al, 2000, Nature Genetics, 24(3), p227-235; Herwig et al., 1999, Genome Res., 9, p1093-1105; Tamayo et al, 1999, Science, PNAS, 96, p2907-2912).

In the clustering method, genes are grouped into functional categories (clusters) based on their expression profile, satisfying two criteria: *homogeneity* - the genes in the same cluster are highly similar in expression to each other; and *separation* - genes in different clusters have low similarity in expression to each other.

Examples of various clustering techniques that have been used for gene expression analysis include hierarchical clustering (Eisen et al., 1998, supra; Alizadeh et al. 2000, supra; Perou et al. 2000, supra; Ross et al, 2000, supra), K-means clustering (Herwig et al., 1999, supra; Tavazoie et al, 1999, Nature Genetics, 22(3), p. 281-285), gene shaving (Hastie et al., 2000, Genome Biology, 1(2), research 0003.1-0003.21), block clustering (Tibshirani et al., 1999, Tech report Univ Stanford.) Plaid model (Lazzeroni, 2002, Stat. Sinica, 12, p61-86), and self-organizing maps (Tamayo et al. 1999, supra). Also, related methods of multivariate statistical analysis, such as those using the singular value decomposition (Alter et al., 2000, PNAS, 97(18),

p10101-10106; Ross et al. 2000, *supra*) or multidimensional scaling can be effective at reducing the dimensions of the objects under study.

However, methods such as cluster analysis and singular value decomposition are purely exploratory and only provide a broad overview of the internal structure present in the data. They are unsupervised approaches in which the available information concerning the nature of the class under investigation is not used in the analysis. Often, the nature of the biological perturbation to which a particular sample has been subjected is known. For example, it is sometimes known whether the sample whose gene expression pattern is being analysed derives from a diseased or healthy individual. In such instances, discriminant analysis can be used for classifying samples into various groups based on their gene expression data.

In such an analysis one builds the classifier by training the data that is capable of discriminating between member and non-members of a given class. The trained classifier can then be used to predict the class of unknown samples. Examples of discrimination methods that have been described in the literature include Support Vector Machines (Brown et al, 2000, *PNAS*, 97, p262-267), Nearest Neighbour (Dudoit et al., 2000, *supra*), Classification trees (Dudoit et al., 2000, *supra*), Voted classification (Dudoit et al., 2000, *supra*), Weighted Gene voting (Golub et al. 1999, *supra*), and Bayesian classification (Keller et al. 2000, Tec report Univ of Washington). Also a technique in which PLS (Partial Least Square) regression analysis is first used to reduce the dimensions in the gene expression data set followed by classification using logistic discriminant analysis and quadratic discriminant analysis (LD and QDA) has recently been described (Nguyen & Rocke, 2002, *Bioinformatics*, 18, p39-50 and

1216-1226).

A challenge that gene expression data poses to classical discriminatory methods is that the number of genes whose expression are being analysed is very large compared to the number of samples being analysed. However in most cases only a small fraction of these genes are informative in discriminant analysis problems. Moreover, there is a danger that the noise from irrelevant genes can mask or distort the information from the informative genes. Several methods have been suggested in literature to identify and select genes that are informative in microarray studies, for example, t-statistics (Dudoit et al, 2002, J. Am. Stat. Ass., 97, p77-87), analysis of variance (Kerr et al., 2000, PNAS, 98, p8961-8965), Neighbourhood analysis (Golub et al, 1999, supra), Ratio of between groups to within groups sum of squares (Dudoit et al., 2002, supra), Non parametric scoring (Park et al., 2002, Pacific Symposium on Biocomputing, p52-63) and Likelihood selection (Keller et al., 2000, supra).

In the methods described herein the gene expression data that has been normalized and standardized is analysed by using Partial Least Squares Regression (PLSR). Although PLSR is primarily a method used for regression analysis of continuous data (see Appendix A), it can also be utilized as a method for model building and discriminant analysis using a dummy response matrix based on a binary coding. The class assignment is based on a simple dichotomous distinction such as breast cancer (class 1) / healthy (class 2), or a multiple distinction based on multiple disease diagnosis such as breast cancer (class 1) / Alzheimer (class 2) / healthy (class 3). The list of diseases for classification can be increased depending upon the samples available corresponding to other diseases or conditions or stages thereof.

PLSR applied as a classification method is referred to as PLS-DA (DA standing for Discriminant analysis). PLS-DA is an extension of the PLSR algorithm in which the Y-matrix is a dummy matrix containing n rows (corresponding to the number of samples) and K columns (corresponding to the number of classes). The Y-matrix is constructed by inserting 1 in the k th column and -1 in all the other columns if the corresponding i th object of X belongs to class k . By regressing Y onto X , classification of a new sample is achieved by selecting the group corresponding to the largest component of the fitted, $\hat{y}(x) = (\hat{y}_1(x), \hat{y}_2(x), \dots, \hat{y}_K(x))$. Thus, in a -1/1 response matrix, a prediction value below 0 means that the sample belongs to the class designated as -1, while a prediction value above 0 implies that the sample belongs to the class designated as 1.

An advantage of PLSR-DA is that the results obtained can be easily represented in the form of two different plots, the score and loading plots. Score plots represent a projection of the samples onto the principal components and shows the distribution of the samples in the classification model and their relationship to one another. Loading plots display correlations between the variables present in the data set.

It is usually recommended to use PLS-DA as a starting point for the classification problem due to its ability to handle collinear data, and the property of PLSR as a dimension reduction technique. Once this purpose has been satisfied, it is possible to use other methods such as Linear discriminant analysis, LDA, that has been shown to be effective in extracting further information, Indahl et al. (1999, Chem. and Intell. Lab. Syst., 49, p19-31). This approach is based on first decomposing the data using PLS-DA, and then using the scores vectors (instead of the original variables) as

input to LDA. Further details on LDA can be found in Duda and Hart (Classification and Scene Analysis, 1973, Wiley, USA).

The next step following model building is of model validation. This step is considered to be amongst the most important aspects of multivariate analysis, and tests the "goodness" of the calibration model which has been built. In this work, a cross validation approach has been used for validation. In this approach, one or a few samples are kept out in each segment while the model is built using a full cross-validation on the basis of the remaining data. The samples left out are then used for prediction/classification. Repeating the simple cross-validation process several times holding different samples out for each cross-validation leads to a so-called double cross-validation procedure. This approach has been shown to work well with a limited amount of data, as is the case in some of the Examples described here. Also, since the cross validation step is repeated several times the dangers of model bias and overfitting are reduced.

Once a calibration model has been built and validated, genes exhibiting an expression pattern that is most relevant for describing the desired information in the model can be selected by techniques described in the prior art for variable selection, as mentioned elsewhere. Variable selection will help in reducing the final model complexity, provide a parsimonious model, and thus lead to a reliable model that can be used for prediction. Moreover, use of fewer genes for the purpose of providing diagnosis will reduce the cost of the diagnostic product. In this way informative probes which would bind to the genes of relevance may be identified.

We have found that after a calibration model has been built, statistical techniques like Jackknife

(Effron, 1982, The Jackknife, the Bootstrap and other resampling plans. Society for Industrial and Applied mathematics, Philadelphia, USA), based on resampling methodology, can be efficiently used to select or confirm significant variables (informative probes).

The approximate uncertainty variance of the PLS regression coefficients B can be estimated by:

$$S^2B = \sum_{m=1}^M ((B-B_m)g)^2$$

where

S^2B = estimated uncertainty variance of B;

B = the regression coefficient at the cross validated rank A using all the N objects;

B_m = the regression coefficient at the rank A using all objects except the object(s) left out in cross validation segment m; and

g = scaling coefficient (here: g=1).

In our approach, Jackknife has been implemented together with cross-validation. For each variable the difference between the B-coefficients B_i in a cross-validated sub-model and B_{tot} for the total model is first calculated. The sum of the squares of the differences is then calculated in all sub-models to obtain an expression of the variance of the B_i estimate for a variable. The significance of the estimate of B_i is calculated using the t-test. Thus, the resulting regression coefficients can be presented with uncertainty limits that correspond to 2 Standard Deviations, and from that significant variables are detected.

No further details as to the implementation or use of this step are provided here since this has been implemented in commercially available software, The

Unscrambler, CAMO ASA, Norway. Also, details on variable selection using Jackknife can be found in Westad & Martens (2000, J. Near Inf. Spectr., 8, p117-124).

The following approach can be used to select informative probes from a gene expression data set:

- a) keep out one unique sample (including its repetitions if present in the data set) per cross validation segment;
- b) build a calibration model (cross validated segment) on the remaining samples using PLSR-DA;
- c) select the significant genes for the model in step b) using the Jackknife criterion;
- d) repeat the above 3 steps until all the unique samples in the data set are kept out once (as described in step a). For example, if 75 unique samples are present in the data set, 75 different calibration models are built resulting in a collection of 75 different sets of significant probes;
- e) select the most significant variables using the frequency of occurrence criterion in the generated sets of significant probes in step d). For example, a set of probes appearing in all sets (100%) are more informative than probes appearing in only 50% of the generated sets in step d).

Once the informative probes for a disease have been selected, a final model is made and validated. The two most commonly used ways of validating the model are cross-validation (CV) and test set validation. In cross-validation, the data is divided into k subsets. The model is then trained k times, each time leaving out one of the subsets from training, but using only the omitted subset to compute error criterion, RMSEP (Root Mean Square Error of Prediction). If k equals the sample size, this is called "leave-one-out" cross-validation. The idea of leaving one or a few samples

out per validation segment is valid only in cases where the covariance between the various experiments is zero. Thus, one sample at-a-time approach can not be justified in situations containing replicates since keeping only one of the replicates out will introduce a systematic bias in our analysis. The correct approach in this case will be to leave out all replicates of the same samples at a time since that would satisfy assumptions of zero covariance between the CV-segments.

The second approach for model validation is to use a separate test-set for validating the calibration model. This requires running a separate set of experiments to be used as a test set. This is the preferred approach given that real test data are available.

The final model is then used to identify a disease, condition or stage thereof in test samples. For this purpose, expression data of selected informative genes is generated from test samples and then the final model is used to determine whether a sample belongs to a diseased or non-diseased class or has a condition or stage thereof.

Thus viewed from a yet further aspect the present invention provides a method of identifying probes useful for diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism, comprising the steps of:

- a) immobilizing a set of oligonucleotide probes, preferably as described hereinbefore, on a solid support;
- b) isolating mRNA from a sample of a normal organism (normal sample), which may optionally be reverse transcribed to cDNA;
- c) isolating mRNA from a sample from an organism, corresponding to the sample and organism of step (b), which is known to have said disease

- or condition or a stage thereof (diseased sample), which may optionally be reverse transcribed to cDNA;
- d) hybridizing the mRNA or cDNA of steps (b) and (c) to said set of immobilized oligonucleotide probes of step (a); and
 - e) assessing the amount of mRNA or cDNA hybridizing to each of said oligonucleotide probes to determine the level of gene expression of genes to which said oligonucleotide probes bind in said normal and diseased samples to generate a gene expression data set for each sample;
 - f) normalizing and standardizing said data set of step (e);
 - g) constructing a calibration model for classification, preferably using the statistical techniques Partial Least Squares Discriminant Analysis (PLS-DA) and Linear Discriminant Analysis (LDA);
 - h) performing JackKnife analysis and identifying those oligonucleotide probes which are required for classification of said disease and normal samples into their respective groups.

Preferably a model for classification purposes is generated by using the data relating to the probes identified according to the above described method. Preferably the sample is as described previously. Preferably the oligonucleotides which are immobilized in step (a) are randomly selected as described below or are the probes as described hereinbefore. Such oligonucleotides may be of considerable length, e.g. if using cDNA (which is encompassed within the scope of the term "oligonucleotide"). The identification of such cDNA molecules as useful probes allows the development

of shorter oligonucleotides which reflect the specificity of the cDNA molecules but are easier to manufacture and manipulate.

The above described model may then be used to generate and analyse data of test samples and thus may be used for the diagnostic methods of the invention. In such methods the data generated from the test sample provides the gene expression data set and this is normalized and standardized as described above. This is then fitted to the calibration model described above to provide classification.

The method described herein can also be used to simultaneously select informative probes for several related and unrelated diseases or conditions. Depending upon which diseases or conditions have been included in the calibration or training set, informative probes can be selected for the said diseases or conditions. The informative probes selected for one disease or condition may or may not be similar to the informative probes selected for another disease or condition of interest. It is the pattern with which the selected genes are expressed in relation to each other during a disease, condition, or stage thereof, that determines whether or not they are informative for the disease, condition or stage thereof.

In other words, informative genes are selected based on how their expression correlates with the expression of other selected informative genes under the influence of responses generated by the disease, condition or stage thereof under investigation. In examples 1 and 2 provided hereinafter, 139 informative probes were selected for breast cancer diagnosis and 182 probes were selected for Alzheimer's disease diagnosis by training the gene expression data set of genes representing 1435 or 758 randomly picked cDNA clones for breast cancer/non breast cancer samples, or

Alzheimer/non-Alzheimer samples, respectively. Among the probes selected for breast cancer and Alzheimer, about 10 probes were informative both for breast cancer and Alzheimer disease diagnosis.

For the purpose of isolating informative probes or identifying several related and unrelated diseases, conditions and stages thereof simultaneously, the gene expression data set must contain the information on how genes are expressed when the subject has a particular disease, condition or stage thereof under investigation.

The data set is generated from a set of healthy or diseased samples, where a particular sample may contain the information of only one disease, condition or stages thereof or may also contain information about multiple diseases, conditions or stages thereof. For example, if the isolation of informative probes for Alzheimer disease, breast cancer and diabetes is sought, whole blood samples can be obtained from an Alzheimer patient who has breast cancer and diabetes. Hence, the method also teaches an efficient experimental design to reduce the number of samples required for isolating informative probes by selecting samples representing more than one disease, condition or stage thereof.

As mentioned previously, in view of the high information content of most transcripts, the identification and selection of informative probes for use in diagnosing, monitoring or identifying a particular disease, condition or stage thereof may be dramatically simplified. Thus the pool of genes from which a selection may be made to identify informative probes may be radically reduced.

Unlike, in prior art technologies where informative probes are selected from a population of thousands of genes that are being expressed in a cell, like in microarray, in the method described herein, the informative probes are selected from a limited number of

randomly obtained genes. For example, from a population of 1435 cDNA clones, randomly picked from a human whole blood cDNA library, we were able to select 139 informative probes for breast cancer diagnosis (see Example 1 and Table 2).

Thus in a preferred aspect of the above mentioned method of identifying probes useful for diagnosing or identifying or monitoring a disease or condition or stage thereof in an organism, said set of oligonucleotides which are immobilized in step (a) are randomly selected from a larger set of oligonucleotides, e.g. from a cDNA library or other oligonucleotide pool, which may be, but is preferably not selected from the set provided herein. Preferably said larger set comprises oligonucleotides which correspond to moderately or highly expressed genes. Thus preferably in methods of the invention, the set of oligonucleotides according to the invention are replaced with a set of oligonucleotides which are randomly selected, e.g. from commercially available oligonucleotide or cDNA libraries.

As referred to herein "random" refers to selection which is not biased based on the extent of information carried by the transcripts in relation to the disease, condition or organism under study, ie. without bias towards their likely utility as informative probes. Whilst a random selection may be made from a pool of transcripts (or related products) which have been biased, e.g. to highly or moderately expressed transcripts, preferably random selection is made from a pool of transcripts not biased or selected by a sequence-based criterion. The larger set may therefore contain oligonucleotides corresponding to highly and moderately expressed genes, or alternatively, may be enriched for those corresponding to the highly and moderately expressed genes.

Random selection from highly and moderately expressed genes can be achieved in a wide variety of ways. A strategy used in this work, but not limiting in itself involves randomly picking a significant number of cDNA clones from a cDNA library constructed from a biological specimen under investigation. Since, in a cDNA library, the cDNA clones corresponding to transcripts present in high or moderate amount are more frequently present than transcripts corresponding to cDNA present in low amount, the former will tend to be picked up more frequently than the latter. A pool of cDNA enriched for those corresponding to highly and moderately expressed genes can be isolated by this approach.

To identify genes that are expressed in high or moderate amount among the isolated population for use in methods of the invention, the information about the relative level of their transcripts in samples of interest can be generated using several prior art techniques. Both non-sequence based methods, such as differential display or RNA fingerprinting, and sequence-based methods such as microarrays or macroarrays can be used for the purpose. Alternatively, specific primer sequences for highly and moderately expressed genes can be designed and methods such as quantitative RT-PCR can be used to determine the levels of highly and moderately expressed genes. Hence, a skilled practitioner may use a variety of techniques which are known in the art for determining the relative level of mRNA in a biological sample.

Especially preferably the sample for the isolation of mRNA in the above described method is as described previously and is preferably not from the site of disease and the cells in said sample are not disease cells and have not contacted disease cells.

The following examples are given by way of

illustration only in which the Figures referred to are as follows:

Figure 1 shows the effect of Direct Standardization (DS) on the Alzheimer data measured in two different series of experiments in which AD denotes Alzheimer's samples and A,B are non-Alzheimer's samples. The samples in both series have been labelled systematically as (xx_7/xx_8), whereas the corrected samples from series 8 (in b,c,d) have been labelled as (xx_c), thus, for example, AD2-7 denotes Alzheimer disease sample number 2 in experiment series 7. The circled spots represent the samples chosen as the transfer samples. The connecting lines in figures b,c,d show the proximity of the replicated samples after applying DS. The dashed lines in figures a,c,d represent the decision boundary separating the classes. These lines have not been drawn on the basis of any statistical criteria, but serve the purpose of visually separating the classes. All the four figures show scores plot (PC1-PC2) from PCA analysis based on (a) non-standardized data, (b) scores plot after direct standardization using 3 transfer samples, (c) scores plot after direct standardization using 4 transfer sample, (d) scores plot after direct standardization using 8 transfer samples;

Figure 2 shows the projection of normal (including benign) and breast cancer samples onto a classification model generated by PLSR-DA using the data of 44 informative genes, in which PC is the principal components and N and C are normal and breast cancer samples, respectively;

Figure 3 shows the projection of individuals with and without Alzheimer's disease onto a classification model generated by PLSR-DA using 182 informative genes;

Figures 4, 6 and 8 show projection plots as Figure 2 in which the classification model is generated using 719, 111 and 345 cDNAs, respectively, wherein PC is the

principal components, N denotes normal and B denotes breast cancer samples;

Figures 5, 7 and 9 show prediction plots based on 3 principal components using the data of 719, 111 and 345 cDNAs, respectively;

Figure 10 shows a projection plot as Figure 3 in which the classification model is generated using 520 cDNAs; and

Figure 11 is the prediction plot corresponding to Figure 10.

Example 1: Diagnosis of Breast Cancer

Methods

Whole blood was obtained from the arms of breast cancer patients and patients with benign tumours (Ullevål and Haukland hospitals in Norway). All of the patients with breast cancer had a malignant tumour of the breast (disease samples). Healthy blood was collected from the above two hospitals, or collected at a Health station at Ås, Norway or at DiaGenic AS, Norway, from the arms of female donors with no reported signs of breast cancer. The blood from healthy individuals or with benign tumours comprise the normal samples. The blood was either collected in tubes containing EDTA and stored immediately at -80°C or was collected in PAXgene tubes and stored for 12-24 hours at room temperature before finally storing them at -80°C before use. Further details of the breast cancer and benign tumour patients from which blood was taken is provided in Table 5.

mRNA was isolated from the blood of the 29 breast cancer patients and 46 normal donors and used to prepare labelled probes by reverse transcribing in the presence of $\alpha^{33}\text{P}$ -dATP. The first strand cDNA of the normal and

diseased samples was bound, separately to 1435 cDNA clones immobilized on a solid support (nylon membrane).

These cDNA clones were randomly picked, without any prior knowledge of their gene sequences, from a cDNA library constructed using whole blood of 550 healthy individuals (Clontech, Palo Alto, USA). These methods were conducted as follows.

For amplification of inserts, bacterial clones were grown in microtiter plates containing 150 μ l LB with 50 μ g/ml carbenicillin, and incubated overnight with agitation at 37°C. To lyse the cells, 5 μ l of each culture were diluted with 50 μ l H₂O and incubated for 12 min. at 95°C. Of this mixture, 2 μ l were subjected to a PCR reaction using 20 pmoles of M13 forward and reverse primer in presence of 1.5 mM MgCl₂. PCR reactions were performed with the following cycling protocol: 4 min. at 95°C, followed by 25 cycles of 1 min. at 94°C, 1 min. at 60°C and 3 min. at 72°C either in a RoboCycler® Temperature Cycler (Stratagene, La Jolla, USA) or DNA Engine Dyad Peltier Thermal Cycler (MJ Research Inc., Waltham, USA). The amplified products were denatured by incubating with NaOH (0.2 M, final concentration) for 30 min. and spotted onto Hybond-N+ membranes (Amersham Pharmacia Biotech, Little Chalfont, UK), using MicroGrid II workstation according to the manufacturer's instructions (BioRobotics Ltd, Cambridge England). The immobilized cDNAs were fixed using a UV cross-linker (Hoefer Scientific Instruments, San Francisco, USA).

In addition to the 1435 cDNAs, the printed arrays also contained controls for assessing background level, consistency and sensitivity of the assay. These were spotted at multiple positions and included controls such as PCR mix (without any insert); positive and negative controls of SpotReport™ 10 array validation system

(Stratagene, La Jolla, USA) and cDNAs corresponding to constitutively expressed genes such as b-actin, g-actin, GAPDH, HOD and cyclophilin. Also, oligonucleotides corresponding to SIX1, b-tubulin, TRP-2, MDM2, Myosin Light C, CD44, Maspin, Laminin, and SRP 19 were included to detect disseminated cancer cells.

The total RNA from blood collected in EDTA tubes was purified using Trizol LS Reagent protocol (Invitrogen/Life Technologies). From blood contained in PAXgene tubes, the total RNA was purified according to the supplier's instructions (PreAnalytiX, Hombrechtikon, Switzerland). Contaminating DNA was removed from the isolated RNA by DNAase I treatment using DNA-free kit (Ambion, Inc. Austin, USA). RNA quality was determined visually by inspecting the integrity of 28S and 18S ribosomal bands following agarose gel electrophoresis. The concentration and purity of extracted RNA was determined by measuring the absorbance at 260 nm and 280 nm. mRNA was isolated from the total RNA using Dynabeads as per the supplier's instructions (DynaL AS, Oslo, Norway).

Labelling and hybridization experiments were performed in batches. The number of samples assayed in each batch varied from six to nine. In the case of samples that were assayed more than once (replicates), aliquots derived from the same mRNA pool were used for probe synthesis. For probe synthesis, aliquots of mRNA corresponding to 4-5 µg of total RNA were mixed together with oligodT_{25NV} (0.5 µg/ml) and mRNA spikes of SpotReport™ 10 array validation system (10 pg; Spike 2, 1 pg), heated to 70°C to remove secondary structures, and then chilled on ice. Probes were prepared in 35µl reaction mixes by reverse transcription in the presence of 50µCi [$\alpha^{33}\text{P}$] dATP, 3.5 µM dATP, 0.6 mM each of dCTP,

dTTP, dGTP, 200 units of SuperScript reverse transcriptase (Invitrogen, LifeTechnologies) and 0.1 M DTT, labelling for 1.5 hr at 42°C. Following synthesis, the enzyme was deactivated for 10 min. at 70°C and mRNA removed by incubating the reaction mix for 20 min. at 37°C in 4 units of Ribo H (Promega, Madison USA). Unincorporated nucleotides were removed using ProbeQuant G 50 Columns (Amersham Biosciences, Piscataway, USA).

Prior to hybridization, the membranes were equilibrated in 4 x SSC for 2 hr at room temperature and prehybridized overnight at 65°C in 10 ml prehybridisation solution (4 x SSC, 0.1 M NaH₂PO₄, 1 mM EDTA, 8% dextran sulphate, 10 x denhardt's solution, 1% SDS). Freshly prepared probes were added to 5 ml of the same prehybridisation solution, and hybridization continued overnight at 65°C. The membranes were washed at 65°C at increasing stringency (2 x 30 min. each in 2 x SSC, 0.1% SDS; 1 x SSC, 0.1% SDS; 0.1 x SSC, 0.1% SDS) to remove unspecific signals.

The amount of labelled first strand cDNA binding to each spot was assessed and quantified using a PhosphorImager to generate a gene expression data set. The data was generated using Phoretix software version 3 (Non Linear Dynamics, England). Background subtraction was performed on the generated data by subtracting the median of the line of pixels around each spot outline from the total intensity obtained from the respective spots.

The background-subtracted data was then normalized and transformed by selecting out 50 lowest and 50 maximum signals from each membrane. This step was to exclude genes that were expressed with a high degree of variance. Since the genes varied from membrane to

membrane, the expression data from 497 genes were removed from the data set. The values for the remaining 938 genes were then normalised by using different approaches such as external controls, dividing each spot by the median intensity of the observed signal in the respective membrane, range normalizing the data from each membrane, and then log transforming the data obtained.

The processed data obtained above was then used to isolate the informative probes by:

- a) keeping one unique sample (including all repetitions of the selected sample) out per cross validation segment;
- b) building a calibration model (cross validated) on the remaining samples using PLSR-DA;
- c) selecting the set of significant genes for the model in step b using the Jackknife criterion;
- d) repeating steps a), b) and c) until all the unique samples were kept out once (hence, in all 75 different calibration models were built (after repeating step b) 75 times), resulting in 75 different sets of significant probes (after repeating step c) 75 times));
- e) selecting significant variables using the frequency of occurrence criterion amongst the 75 different sets of significant probes.

The selected informative probes based on occurrence criterion were used to construct a classification model. The result of the classification model based on probes appearing in at least 90% of the generated sets after the step of isolating informative probes as described above is shown in Figure 2 in which it is seen that the expression pattern of these genes was able to classify most women with breast cancer and women with no breast cancer into distinct groups. In this figure PC1 and PC2

indicate the two principal components statistically derived from the data which best define the systemic variability present in the data. This allows each sample, and the data from each of the informative probes to which the sample's labelled first strand cDNA was bound, to be represented on the classification model as a single point which is a projection of the sample onto the principal components - the score plot.

The ability of the generated model, based on isolated informative probes, to predict future samples was determined by the double cross-validation approach. The performance of the diagnostic test for breast cancer based on the occurrence criterion is presented in Table 6.

Correct prediction of most breast cancer cells was achieved. These included all three samples obtained from women with ductal carcinoma in situ (DCIS), 11/15 samples obtained from women with stage I breast cancer, all five samples obtained from women with stage II breast cancer, and one of two samples obtained from women with stage III breast cancer. Interestingly, two correctly predicted stage I samples were obtained from women having a tumour size of <5 mm in diameter.

The model also correctly predicted the class of most non-cancer samples (41/46), including those that were obtained from women with non-cancerous breast abnormalities.

Confirmation that the gene transcripts are not from cells which are disseminated disease cells has been confirmed by several lines of evidences. Firstly, the informative genes were expressed constitutively at high or moderate levels in blood cells of women irrespective

of whether they had cancer or not. Secondly, in the assay described in this Example, in order to identify transcripts, at least 720 disseminated cells in blood samples would be required. Since, the average number of disseminated cells present in blood during different stages of breast cancer is much lower (organ confined breast cancer, 0.8 cells per ml; invasive breast cancer spread to lymph nodes only, 2.4 cells per ml; and metastatic breast cancer, 6 cells per ml; SD>100%) (29), we believe that the signals being detected originated from peripheral blood cells and could not have originated from disseminated cells. Thirdly, we were not able to detect any signal from the eight cancer markers known to have elevated expression in malignant cancer cells, including cancer cells that are disseminated in the blood.

Example 2: Diagnosis of Alzheimer's disease

Similar experiments were conducted with samples from Alzheimer's patients. In this method 7 patients diagnosed with Alzheimer's Disease at the Memory Clinic at Ullevål University Hospital were used in the trial. The patients were confirmed as having Alzheimer's disease based on the following criteria:

- * A standardized interview with a care-giver using IQCODE, an ADL scale and a scale measuring behaviour of the patient (Green scale).
- * Neuropsychological evaluation using MMSE, Clock drawing test, Trailmaking test A and B (TMT A and B), Kendrick object learning test (visual memory test), part of the Wechsler battery and Benton test.
- * A psychiatric evaluation using scales for detection of depression, MADRS for interviewing the patient and Cornell scale for interviewing the care-giver.

- * A physical examination.
- * Laboratory tests of blood samples to rule out other diseases.
- * CT scan of the brain.
- * SPECT of the brain.

The mean age of the patients was 72.3 with an age range of 69-76. The mean MMSE score was 22.0 (the maximum score attainable being 30).

Six age-matched individuals without diagnosed Alzheimer's disease were used as a control. All had been tested with MMSE and had a minimum score of 28 (mean: 28.4). The mean age of the normal control group was 73.0 and the age range 66-81. A sample from a 16-year old individual, with a consequent minimal chance of having Alzheimer's disease, was also included as an additional control.

Using the methods described above (except that hybridization to 758 rather than 1435 cDNA clones was performed), informative probes were selected based on occurrence criterion and used to construct a classification model. The results of the classification model based on probes appearing at least once in the generated sets after the method to isolate informative probes as described above is shown in Figure 3 in which it will be seen that the expression pattern of these genes was able to classify individuals with or without Alzheimer's disease into distinct groups. In this Figure PC1 and PC2 indicate the 2 principal components statistically derived from the data which define the systematic variability present in the data. This allows each sample, and the data from each of the informative probes to which the samples' cDNA was bound, to be represented on the classification model as a single

point which is a projection of the sample onto the principal components - the score plot.

The ability of the generated model, based on isolated informative probes, to predict future samples was determined by the double cross-validation. The performance of the diagnostic test for Alzheimer's disease is presented in Table 7.

Appendix A

Partial Least Squares regression (PLSR)

Let a multivariate regression model be defined as:

$$Y = XB + F$$

where

X a $N \times P$ matrix with N predictor variables (genes);
 Y ($N \times J$) being the J predicted variables. In our case Y represents a matrix containing dummy variables;
 B is a matrix of regression coefficients; and
 F is a $N \times J$ matrix of residuals.

The structure of the PLSR model can be written as:

$$X = TP^T + E_A, \text{ and}$$

$$Y = TQ^T + F_A, \text{ where}$$

where

T ($N \times A$) is a matrix of score vectors which are linear combinations of the x -variables;
 P ($P \times A$) is a matrix with the x -loading vectors p_a as columns;
 Q ($J \times A$) is a matrix with the y -loading vectors q_a as columns;
 E_a ($N \times P$) is the matrix for X after A factors; and
 F_a ($N \times J$) is the matrix for Y after A factors.

The criterion in PLSR is to maximize the explained covariance of $[X, Y]$. This is achieved by the loading weights vector w_{a+1} , which is the first eigenvector of $E_a^T F_a F_a^T E_a$ (E_a and F_a are the deflated X and Y after a factors or PLS components).

The regression coefficients are given by:

$$B = W(P^T W)^{-1} Q^T$$

A PLSR model with full rank, i.e. maximum number of components, is equivalent to the MLR solutions. Further details on PLSR can be found in Martens & Naes, 1989, Multivariate Calibration, John Wiley & Sons, Inc., USA and Kowalski & Seasholtz, 1991, *supra*.

Example 3: Validation of Example 1, diagnosis of breast cancer

The results in Example 1 were validated by using the informative probes identified in Example 1 on new breast cancer and control samples.

Methods

The methods, essentially as described in Example 1, were used. Blood was taken from patients as described in Table 8. However, blood was collected in PAXgene tubes and the first strand labelled cDNAs were hybridized to 719 cDNAs spotted on nylon membranes along with other controls as described in Example 1. After background subtraction using control spots, the data of each membrane was normalized using the inter quantile range.

The data was analysed as described in Example 1 and the model validated by cross validation.

The 719 cDNAs which were spotted are a subset of the cDNAs spotted in Example 1 and include 111 cDNAs described in Table 2 and which were found to be informative in Example 1.

Results

The results are shown in Figures 4 to 9. Figures 4, 6 and 8 are projection plots similar to Figure 2 and show the projection of normal and breast cancer patients' samples onto a classification model generated using all 719 cDNA. Figure 6 is similar but uses a classification model generated with the 111 probes common to Example 1.

Figure 8 uses the 345 sequences of the 719 for which sequence information is provided herein. In each case classification of normal and breast cancer groups was possible. Figures 5, 7 and 9 show prediction plots which reflect the ability of the generated models to

correctly diagnose breast cancer. In the 3 prediction plots shown, the disease samples appear on the x axis at +1 and the non-disease samples appear at -1. The y axis represents the predicted class membership. During prediction, if the prediction is correct, disease samples should fall above zero and non-disease samples should fall below zero. In each case almost all samples are correctly predicted.

Example 4: Validation of Example 2, diagnosis of Alzheimers

The results in Example 2 were validated by using the informative probes identified in Example 2 on new Alzheimer's patient samples.

Methods

The methods, essentially as described in Example 2, were used. Twelve female patients diagnosed with Alzheimer's disease at the Memory Clinic at Ullevål University Hospital who were confirmed as having Alzheimer's disease based on the criteria of Example 2 were used in the trial. The mean age of the patients was 72.3 with an age range of 66-83. The mean MMSE score was 22.0 (the maximum score attainable being 30).

Sixteen age-matched female individuals without diagnosed Alzheimer's disease were used as the normal control group. All had been tested with MMSE and had a minimum score of 29. The mean age of the normal control group was 74.0 and the age range 66-86.

After transfer of the blood to PAXgene tubes, total mRNA was isolated from the blood of the Alzheimer's disease and from the control group donors according to the manufacturers's instructions (PreAnalytiX,

Hombrechtikon, Switzerland). The isolated mRNA was labelled during reverse transcription in the presence of $\alpha^{33}\text{P}$ -dATP, yielding a labelled first strand cDNA. Hybridization was performed as described previously onto 730 cDNA clones picked from a cDNA library from whole blood of 550 healthy individuals without knowledge of the gene sequence of the random cDNA clones.

Results

The results are shown in Figures 10 and 11. Figure 10 is a projection plot generated using 520 probes which have been sequenced. Figure 11 is a prediction plot and shows correct prediction of almost all samples.

Table 1a

List of probes informative for disease diagnosis

	Clone ID	Sequence ID	No. of nucleotides	SEQ ID NO: in <u>sequence listing</u>
1	I-01	-	-	
2	I-02	-	-	
3	I-13	-	-	
4	I-21	-	-	
<u>51</u>	I-24	308	373	<u>11</u>
62	I-28	310	564	<u>13</u>
73	I-30	1180	622	<u>398</u>
84	I-34	313	554	<u>15</u>
9	I-37	-	-	
10	I-42	-	-	
11	I-62	-	-	
125	I-54	1181	156	<u>399</u>
136	I-58	326	554	<u>24</u>
14	I-71	-	-	
15	I-72	-	-	
16	I-86	-	-	
17	I-95	-	-	
187	II-03	361	622	<u>34</u>
198	II-05	363	628	<u>35</u>
209	II-06	364	528	<u>36</u>
2110	II-10	368	329	<u>39</u>
2211	II-24	381	534	<u>47</u>
2312	II-25	382	444	<u>48</u>
2413	II-26	383	566	<u>49</u>
2514	II-33	390	523	<u>55</u>
2615	II-34	391	566	<u>56</u>
2716	II-41	397	534	<u>60</u>
				61

28 <u>17</u>	II-42	398	512	
29	II-47	-	-	
30 <u>18</u>	II-57	411	505	<u>73</u>
31 <u>19</u>	II-61	415	596	<u>77</u>
32 <u>20</u>	II-69	423	387	<u>85</u>
33 <u>21</u>	II-70	424	420	<u>86</u>
34 <u>22</u>	II-75	429	535	<u>91</u>
35	II-83	-	-	
36 <u>23</u>	II-84	438	577	<u>99</u>
37 <u>24</u>	II-87	441	552	<u>100</u>
38 <u>25</u>	II-88	442	606	<u>101</u>
39	II-90	-	-	
40 <u>26</u>	II-94	448	329	<u>104</u>
41 <u>27</u>	III-02	453	747	<u>107</u>
42	III-05	-	-	
43 <u>28</u>	III-06	458	682	<u>109</u>
44 <u>29</u>	III-08	460	536	<u>111</u>
45	III-10	-	-	
46 <u>30</u>	III-13	464	615	<u>115</u>
47	III-15	-	-	
48	III-17	-	-	
49 <u>31</u>	III-20	483	479	<u>401</u>
50 <u>32</u>	III-23	473	694	<u>119</u>
51 <u>33</u>	III-26	476	476	<u>122</u>
52 <u>34</u>	III-35	485	551	<u>130</u>
53 <u>35</u>	III-39	487	224	<u>131</u>
54 <u>36</u>	III-40	488	349	<u>132</u>
55 <u>37</u>	III-43	490	382	<u>500</u>
56 <u>38</u>	III-44	491	382	<u>134</u>
57 <u>39</u>	III-53	500	390	<u>142</u>
58 <u>40</u>	III-56	503	109	<u>144</u>
59 <u>41</u>	III-57	504	374	<u>145</u>
60	III-60	556	325	

61	III-60	-	-	
62 <u>42</u>	III-61	507	521	<u>148</u>
63 <u>43</u>	III-63	509	575	<u>150</u>
64	III-68	-	-	
65 <u>44</u>	III-74	518	502	<u>155</u>
66 <u>45</u>	III-80	523	585	<u>158</u>
67	III-82	-	-	
68 <u>46</u>	III-85	526	516	<u>161</u>
69 <u>47</u>	III-89	530	660	<u>165</u>
70	III-92	-	-	
71	III-96	-	-	
72 <u>48</u>	IV-14	684	545	<u>275</u>
73 <u>49</u>	IV-15	1185	628	<u>402</u>
74	IV-23	-	-	
76 <u>50</u>	IV-26	1186	494	<u>403</u>
75	IV-26	-	-	
77	IV-29	-	-	
78 <u>51</u>	IV-31	687	268	<u>278</u>
79 <u>52</u>	IV-32	688	569	<u>279</u>
80	IV-34	-	-	
81	IV-35	-	-	
82	IV-41	-	-	
83	IV-45	-	-	
84 <u>53</u>	IV-53	61	362	<u>498</u>
85	IV-62	-	-	
86 <u>54</u>	IV-69	192	286	<u>4</u>
87 <u>55</u>	IV-80	701	579	<u>291</u>
88	IV-82	-	-	
89	IV-93	-	-	
90 <u>56</u>	IX-10	736	641	<u>314</u>
91	IX-12	-	-	
92 <u>57</u>	IX-38	757	583	<u>317</u>
93 <u>58</u>	IX-39	758	424	<u>318</u>

94	IX-42	-	-	
95 <u>59</u>	IX-48	764	626	<u>319</u>
96 <u>60</u>	IX-77	785	556	<u>325</u>
97	V-01	-	-	
98	V-02	-	-	
99 <u>61</u>	V-03	706	496	<u>296</u>
100 <u>62</u>	V-04	707	397	<u>297</u>
101	V-06	-	-	
102 <u>63</u>	V-07	708	293	<u>298</u>
103 <u>64</u>	V-11	1188	599	<u>404</u>
104 <u>65</u>	V-12	711	498	<u>301</u>
105	V-15	-	-	
106	V-17	-	-	
107	V-21	-	-	
108	V-25	-	-	
109	V-32	-	-	
110	V-35	-	-	
111	V-39	-	-	
112	V-42	-	-	
113	V-43	-	-	
114	V-47	-	-	
115	V-49	-	-	
116	V-52	-	-	
117	V-54	-	-	
118 <u>66</u>	V-55	77	421	<u>499</u>
119	V-58	-	-	
120	V-59	-	-	
121	V-65	-	-	
122	V-68	-	-	
123	V-71	-	-	
124	V-75	-	-	
125	V-79	-	-	
126 <u>67</u>	V-80	726	260	<u>311</u>

127	V-90	-	-	
128	V-91	-	-	
129	V-92	-	-	
130	V-94	-	-	
131	VI-02	-	-	
132 <u>68</u>	VI-04	865	122	<u>339</u>
133 <u>69</u>	VI-07	93	405	<u>1</u>
134	VI-09	-	-	
135	VI-10	-	-	
136 <u>70</u>	VI-12	869	667	<u>341</u>
137 <u>71</u>	VI-14	871	642	<u>343</u>
138	VI-17	-	-	
139 <u>72</u>	VI-20	876	115	<u>346</u>
140	VI-21	-	-	
141 <u>73</u>	VI-23	878	634	<u>347</u>
142	VI-34	-	-	
143	VI-41	-	-	
144	VI-42	-	-	
145	VI-43	-	-	
146	VI-44	-	-	
147 <u>74</u>	VI-48	891	626	<u>355</u>
148	VI-49	-	-	
149 <u>75</u>	VI-50	893	585	<u>356</u>
150	VI-52	-	-	
151 <u>76</u>	VI-53	895	560	<u>357</u>
152 <u>77</u>	VI-55	897	509	<u>359</u>
153	VI-65	-	-	
154 <u>78</u>	VI-70	108	550	<u>2</u>
155	VI-71	-	-	
156	VI-72	-	-	
157 <u>79</u>	VI-74	905	655	<u>365</u>
158 <u>80</u>	VI-76	907	582	<u>367</u>
159	VI-78	-	-	

160	VI-79	-	-	
161	VI-84	-	-	
162 <u>81</u>	VI-87	911	595	<u>370</u>
163 <u>82</u>	VI-88	912	651	<u>371</u>
164	VI-90	-	-	
165	VI-93	-	-	
166 <u>83</u>	VI-95	915	230	<u>374</u>
167	VI-96	-	-	
168	VII-02	-	-	
169 <u>84</u>	VII-03	1196	412	<u>411</u>
170	VII-06	-	-	
171	VII-10	-	-	
172	VII-11	-	-	
173 <u>85</u>	VII-15	1199	439	<u>414</u>
174 <u>86</u>	VII-19	562	580	<u>171</u>
175 <u>87</u>	VII-21	564	671	<u>173</u>
176	VII-25	-	-	
177 <u>88</u>	VII-32	571	457	<u>179</u>
178 <u>89</u>	VII-36	575	209	<u>182</u>
179 <u>90</u>	VII-39	576	541	<u>183</u>
180 <u>91</u>	VII-42	579	502	<u>186</u>
181 <u>92</u>	VII-43	580	316	<u>187</u>
182 <u>93</u>	VII-46	583	631	<u>190</u>
183 <u>94</u>	VII-47	1200	526	<u>415</u>
184 <u>95</u>	VII-48	1201	613	<u>416</u>
185 <u>96</u>	VII-59	593	565	<u>199</u>
186	VII-60	-	-	
187 <u>97</u>	VII-63	595	98	<u>201</u>
188 <u>98</u>	VII-66	598	362	<u>204</u>
189	VII-67	-	-	
190 <u>99</u>	VII-72	600	595	<u>206</u>
191 <u>100</u>	VII-73	601	522	<u>207</u>
192	VII-75	-	-	
				209

193 <u>101</u>	VII-76	603	624	
194 <u>102</u>	VII-77	1203	692	<u>418</u>
195 <u>103</u>	VII-80	605	338	<u>210</u>
196 <u>104</u>	VII-81	606	556	<u>211</u>
197	VII-83	-	-	
198	VII-86	-	-	
199	VII-88	-	-	
200 <u>105</u>	VII-90	612	576	<u>216</u>
201 <u>106</u>	VII-91	613	341	<u>217</u>
202 <u>107</u>	VII-93	615	379	<u>219</u>
203	VIII-01	-	-	
204	VIII-02	-	-	
205	VIII-03	-	-	
206	VIII-06	-	-	
207 <u>108</u>	VIII-09	619	598	<u>221</u>
208	VIII-10	-	-	
209	VIII-15	-	-	
210 <u>109</u>	VIII-20	629	419	<u>229</u>
211	VIII-22	-	-	
212	VIII-26	-	-	
213 <u>110</u>	VIII-28	634	511	<u>235</u>
214 <u>111</u>	VIII-29	635	592	<u>236</u>
215 <u>112</u>	VIII-30	636	572	<u>237</u>
216 <u>113</u>	VIII-31	637	482	<u>238</u>
217 <u>114</u>	VIII-32	638	545	<u>239</u>
218 <u>115</u>	VIII-33	639	624	<u>240</u>
219	VIII-39	-	-	
220 <u>116</u>	VIII-41	645	649	<u>245</u>
221 <u>117</u>	VIII-42	646	600	<u>246</u>
222	VIII-44	-	-	
223 <u>118</u>	VIII-46	649	425	<u>249</u>
224 <u>119</u>	VIII-48	651	251	<u>251</u>
225	VIII-58	-	-	
				261

226 <u>120</u>	VIII-64	663	627	
227	VIII-65	-	-	
228 <u>121</u>	VIII-66	665	345	<u>262</u>
229 <u>122</u>	VIII-67	666	252	<u>263</u>
230	VIII-74	-	-	
231 <u>123</u>	VIII-76	675	691	<u>270</u>
232	VIII-78	-	-	
233	VIII-82	-	-	
234	VIII-83	-	-	
235	VIII-85	-	-	
236	VIII-87	-	-	
237	VIII-91	-	-	
238	VIII-92	-	-	
239	VIII-93	-	-	
240	VIII-95	-	-	
241	X-04	-	-	
242 <u>124</u>	X-07	808	641	<u>328</u>
243 <u>125</u>	X-15	814	132	<u>329</u>
244 <u>126</u>	X-29	821	370	<u>331</u>
245	X-34	-	-	
246	X-35	-	-	
247 <u>127</u>	X-54	837	603	<u>334</u>
248 <u>128</u>	X-56	839	71	<u>335</u>
249 <u>129</u>	X-68	1207	642	<u>421</u>
250 <u>130</u>	X-72	849	622	<u>336</u>
251 <u>131</u>	X-94	860	601	<u>337</u>
252	XI-07	-	-	
253 <u>132</u>	XI-13	1209	620	<u>423</u>
254	XI-50	-	-	
255	XI-58	-	-	
256 <u>133</u>	XI-81	1212	374	<u>426</u>
257 <u>134</u>	XII-07	1213	567	<u>427</u>
258	XII-17	-	-	

259	XII-26	-	-	
260	XII-27	-	-	
261	XII-31	-	-	
262	XII-32	-	-	
263 <u>135</u>	XII-35	1214	620	<u>428</u>
264	XII-36	-	-	
265	XII-52	-	-	
266 <u>136</u>	XII-59	1216	484	<u>430</u>
267 <u>137</u>	XIII-19	1219	559	<u>433</u>
268	XIII-29	-	-	
269 <u>138</u>	XIII-52	939	513	<u>378</u>
270	XIII-62	-	-	
271	XIII-84	-	-	
272 <u>139</u>	XIII-92	1221	741	<u>435</u>
273	XV-18	-	-	
274 <u>140</u>	XV-22	-	-	<u>388</u>
275	XV-24	-	-	
276 <u>141</u>	XV-25	1224	485	<u>436</u>
277	XV-28	-	-	
278	XV-34	-	-	
279	XV-42	-	-	
280	XV-68	-	-	
281	XV-74	-	-	
282	XV-93	-	-	
283	XV-94	-	-	
284	XV-96	-	-	
285 <u>142</u>	XVI-36	1056	435	<u>382</u>
286 <u>143</u>	XVI-53	1230	741	<u>439</u>
287	XVI-59	-	-	
288 <u>144</u>	XVI-66	1074	689	<u>384</u>
289 <u>145</u>	XVI-76	1083	198	<u>386</u>
290 <u>146</u>	XVI-77	1084	198	<u>387</u>
291	XVII-07	-	-	

292	XVII-08	-	-	
293	XVII-17	-	-	
294	XVII-28	-	-	
295	XVII-29	-	-	
296 <u>147</u>	XVII-31	1139	503	<u>392</u>
297	XVII-36	-	-	
298	XVII-39	-	-	
299 <u>148</u>	XVII-40	1231	203	<u>440</u>
300 <u>149</u>	XVII-48	1148	587	<u>393</u>
301	XVII-55	-	-	
302	XVII-58	-	-	
303	XVII-67	-	-	
304	XVII-72	-	-	
305 <u>150</u>	XVII-76	1160	650	<u>394</u>
306	XVII-82	-	-	
307 <u>151</u>	XVII-87	1165	502	<u>395</u>
308 <u>152</u>	XVII-95	1172	648	<u>396</u>

Table 1b

List of sequences of probes informative for disease diagnosis
Please see the note at the bottom

Clone ID	Sequence ID NO. in Sequence Listing
I-09	290
I-10	2996
I-13	1331444
I-14	1179397
I-15	3007
I-16	301
I-17	3028
I-19	3049
I-20	305
I-22	30610
I-23	307
I-24	30811
I-25	30912
I-28	31013
I-30	1180398
I-31	31114
I-32	312
I-34	31315
I-37	1440482
I-38	31416
I-39	31517
I-40	31618
I-42	1332445
I-44	317
I-45	318
I-46	319
I-47	320
I-48	32119
I-49	32220
I-53	32321
I-54	1181399
I-56	32422
I-57	32523
I-58	32624
I-60	32725
I-64	32826
I-67	33027
I-69	33128
I-71	332
I-72	333
I-73	334
I-77	33529
I-79	336
I-80	33730
I-81	33831
I-82	33932
I-86	1334447
I-88	1182400
I-95	1337448
II-02	36033
II-03	36134
II-05	36335
II-06	36436
II-07	36537
II-08	36638
II-09	367
II-10	36839
II-11	36940
II-12	37041
II-13	37142
II-14	372
II-15	37343
II-16	37444

II-17	375
II-18	376
II-20	377
II-21	37845
II-22	379
II-23	38046
II-24	38147
II-25	38248
II-26	38349
II-27	38450
II-28	385
II-29	38651
II-30	38752
II-31	38853
II-32	38954
II-33	39055
II-34	39156
II-35	392
II-37	393
II-38	39457
II-39	39558
II-40	39659
II-41	39760
II-42	39861
II-43	39962
II-44	40063
II-46	40164
II-47	40265
II-48	40366
II-49	404
II-50	40567
II-52	40668
II-53	40769
II-54	40870
II-55	40971
II-56	41072
II-57	41173
II-58	41274
II-59	41375
II-60	41476
II-61	41577
II-62	41678
II-63	41779
II-64	41880
II-65	41981
II-66	42082
II-67	42183
II-68	42284
II-69	42385
II-70	42486
II-71	42587
II-72	42688
II-73	42789
II-74	42890
II-75	42991
II-76	43092
II-77	43193
II-78	43294
II-79	43395
II-80	43496
II-81	43597
II-82	43698
II-83	437
II-84	43899
II-85	439
II-86	440
II-87	441100
II-88	442101
II-89	443
II-90	444
II-91	445
II-92	446102
II-93	447103

II-94	448104
II-95	449
II-96	450105
III-01	452106
III-02	453107
III-03	454108
III-04	455
III-05	457
III-06	458109
III-07	459110
III-08	460111
III-09	461112
III-11	462113
III-12	463114
III-13	464115
III-14	465
III-15	466
III-16	467
III-17	468
III-18	469116
III-19	470
III-20	1183401
III-21	471117
III-22	472118
III-23	473119
III-24	474120
III-25	475121
III-26	476122
III-27	477123
III-28	478124
III-29	479125
III-31	481126
III-32	482127
III-33	483128
III-34	484129
III-35	485130
III-37	486
III-39	487131
III-40	488132
III-42	489133
III-43	490500
III-44	491134
III-45	492135
III-46	493136
III-47	494137
III-48	495138
III-49	496139
III-50	497140
III-51	498
III-52	499141
III-53	500142
III-54	501
III-55	502143
III-56	503144
III-57	504145
III-58	505146
III-59	506147
III-61	507148
III-62	508149
III-63	509150
III-64	510151
III-65	511
III-66	512152
III-67	513153
III-69	514
III-70	515154
III-71	516
III-73	517
III-74	518155
III-76	519156
III-77	520
III-78	521157
III-79	522

III-80	523158
III-81	524159
III-82	1340451
III-83	525160
III-85	526161
III-86	527162
III-87	528
III-88	529163 & 164
III-89	530165
III-91	531
III-92	1351452
III-93	532166
III-94	533167
III-95	534168
III-96	535
IV-02	681
IV-04	682273
IV-13	683274
IV-14	684275
IV-15	1185402
IV-17	685276
IV-23	1353454
IV-26	1186403
IV-28	686277
IV-31	687278
IV-32	688279
IV-35	1355455
IV-37	689280
IV-38	689280
IV-40	690281
IV-42	691282
IV-43	1239441
IV-44	692283
IV-47	693284
IV-53	61498
IV-55	694285
IV-56	695
IV-61	696286
IV-64	697287
IV-65	698288
IV-69	1924
IV-72	699289
IV-73	700290
IV-80	701291
IV-82	196
IV-85	702292
IV-93	703293
IV-95	704294
IV-96	705295
IX-10	736314
IX-12	738
IX-13	739315
IX-24	747316
IX-38	757317
IX-39	758318
IX-48	764319
IX-50	766320
IX-56	768321
IX-62	773322
IX-65	776323
IX-72	782324
IX-77	785325
IX-91	796326
IX-96	801327
V-01	1361458
V-03	706296
V-04	707297
V-07	708298
V-08	709299
V-09	710300
V-11	1188404
V1-16	873344
V1-19	875345

V-12	711301
V-17	1364459
V-18	712
V-20	713302
V-24	714303
V-25	1365460
V-28	1109405
V-35	1366461
V-37	716
V-38	1190406
V-39	1109389
V-40	717304
V-41	718305
V-47	1368463
V-48	719306
V-49	1369464
V-55	77499
V-57	720307
V-58	1370465
V-61	721308
V-64	722309
V-65	723
V-68	1448484
V-71	1495496
V-74	724310
V-75	1372467
V-80	726311
V-81	727312
V-87	728313
V-90	1374468
VI-02	340
VI-03	341
VI-04	342
VI-06	343
VI-07	344
VI-08	345
VI-09	346
VI-11	347
VI-12	869341
VI-13	870342
VI-14	871343
VI-16	873344
VI-18	348
VI-19	349
VI-20	350
VI-21	351
VI-22	352
VI-23	878347
VI-24	879348
VI-25	353
VI-26	354
VI-27	355
VI-31	356
VI-32	885351
VI-33	357
VI-35	358
VI-39	887352
VI-43	1382471
VI-44	1193409
VI-45	889353
VI-48	359
VI-49	892501
VI-50	893356
VI-53	895357
VI-55	897359
VI-58	899361
VI-66	903363
VI-67	904364
VI-70	1082
VI-71	1387472
VI-74	905365
VI-75	906366
VI-76	907367

VI-77	4103
VI-79	1389473
VI-80	908368
VI-85	910369
VI-87	911370
VI-88	912371
VI-90	1390474
VI-93	1391475
VI-95	915374
VI-96	1392476
VII-02	547
VII-03	548
VII-04	549
VII-05	550
VII-06	551
VII-07	552
VII-08	553
VII-09	554
VII-10	555
VII-11	556
VII-12	557
VII-14	558
VII-15	559
VII-17	560169
VII-18	561170
VII-19	562171
VII-20	563172
VII-21	564173
VII-22	565174
VII-23	566175
VII-24	567176
VII-25	1397480
VII-26	2505
VII-27	568177
VII-28	569
VII-29	570178
VII-32	571179
VII-33	572180
VII-34	573
VII-35	574181
VII-36	575182
VII-39	576183
VII-40	577184
VII-41	578185
VII-42	579186
VII-43	580187
VII-44	581188
VII-45	582189
VII-46	583190
VII-47	1209415
VII-48	584
VII-49	585191
VII-50	586192
VII-52	587193
VII-53	588194
VII-54	589195
VII-55	590196
VII-57	591197
VII-58	592198
VII-59	593199
VII-62	594200
VII-63	595201
VII-64	596202
VII-65	597203
VII-66	598204
VII-67	1399481
VII-71	599205
VII-72	600206
VII-73	601207
VII-74	602208
VII-76	603209
VII-77	604
VII-80	605210

VII-81	606211
VII-82	607212
VII-83	608
VII-84	609213
VII-86	1453487
VII-87	610214
VII-89	611215
VII-90	612216
VII-91	613217
VII-92	614218
VII-93	615219
VII-94	616
VII-96	617220
VIII-09	618221
VIII-10	619222
VIII-11	620
VIII-12	621223
VIII-13	622224
VIII-15	623
VIII-16	624225
VIII-17	625226
VIII-18	626227
VIII-19	627228
VIII-20	628229
VIII-21	629230
VIII-22	1455
VIII-23	630231
VIII-24	631232
VIII-25	632233
VIII-26	1456489
VIII-27	633234
VIII-28	634235
VIII-29	635236
VIII-30	636237
VIII-31	637238
VIII-32	638239
VIII-33	639240
VIII-34	640
VIII-36	641241
VIII-37	642242
VIII-38	643243
VIII-40	644244
VIII-41	645245
VIII-42	646246
VIII-43	647247
VIII-45	648248
VIII-46	649249
VIII-47	650250
VIII-48	651251
VIII-50	652252
VIII-51	653253
VIII-53	654254
VIII-54	655255
VIII-55	656256
VIII-56	657257
VIII-57	658258
VIII-58	659
VIII-59	660259
VIII-60	661260
VIII-61	662
VIII-64	663261
VIII-65	664
VIII-66	665262
VIII-67	666263
VIII-68	667
VIII-69	668
VIII-70	669264
VIII-71	670265
VIII-72	671266
VIII-73	672267
VIII-74	673268
VIII-75	674269
VIII-76	675270

VIII-77	676271
VIII-78	677
VIII-79	678
VIII-80	679272
X-07	808328
X-15	814329
X-20	817330
X-29	821331
X-34	825332
X-46	833333
X-54	837334
X-56	839335
X-68	1207421
X-72	849336
X-73	1208422
X-94	860337
XI-13	1209423
XI-37	1460490
XI-43	1210424
XI-67	1211425
XI-81	1212426
XII-07	1213427
XII-35	1214428
XII-36	1215429
XII-59	1216430
XII-65	1028381
XII-92	1217431
XIII-03	917375
XIII-04	1218432
XIII-19	1219433
XIII-24	926376
XIII-51	938377
XIII-52	939378
XIII-67	947379
XIII-69	949380
XIII-88	1220434
XIII-92	1221435
XV-22	1099388
XV-24	1101
XV-25	1224436
XV-42	1108
XV-62	1226437
XV-64	1118390
XV-84	1125391
XVI-19	1228438
XVI-36	1056382
XVI-53	1230439
XVI-60	1071383
XVI-66	1074384
XVI-74	1081385
XVI-76	1083386
XVI-77	1084387
XVII-31	1139392
XVII-40	1231440
XVII-48	1148393
XVII-76	1160394
XVII-87	1165395
XVII-95	1172396

Note

~~Sequences not available for sequence IDs in Table 1, and corresponding sequence IDs in Table 2 and 4.~~

~~298, 301, 305, 307, 312, 317, 318, 319, 320, 332, 333, 334, 336, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 367, 372, 375, 376, 377, 379, 385, 392, 393, 404, 437, 439, 440, 443, 444, 445, 449, 455, 457, 465, 466, 467, 468, 470, 486, 498, 501, 511, 514, 516, 517, 520, 522, 528, 531, 535, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 573, 584, 604, 608, 616, 620, 623, 640, 659, 662, 664, 667, 668, 673, 677, 678, 679, 681, 695, 702,~~

~~712, 716, 825, 886, 894, 902, 909, 916, 1101, 1108, 1109, 1177, 1187, 1193, 1204,~~
~~1220, 1239, 1255, 1256, 1342, 1347, 1354, 1357, 1362, 1363, 1364, 1373, 1375, 1379,~~
~~1403, 1404, 1405, 1406, 1413~~

Table 2a

List of informative probes for diagnosis of breast cancer

Clone ID	Sequence- ID SEQ ID NO. in Sequence Listing	Clone ID	Sequence- ID SEQ ID NO. in Sequence Listing
I-24	308 11	VI-72	-
I-28	310 13	VI-78	-
I-30	1180 398	VI-84	-
I-52	-	VII-03	1196 411
I-54	1181 399	VII-15	1199 414
II-41	397 60	VII-32	571 179
II-70	424 86	VII-39	576 183
II-87	441 100	VII-47	1200 415
III-06	458 109	VII-48	1201 416
III-20	1183 401	VII-60	-
III-40	488 132	VII-73	601 207
III-57	504 145	VII-77	1203 418
III-60	-	VII-90	612 216
III-61	507 148	VIII-20	628 229
III-89	530 165	VIII-29	635 236
IV-14	684 275	VIII-30	636 237
IV-15	1185 402	VIII-31	637 238
IV-26	1186 403	VIII-39	-
IV-32	688 279	VIII-44	-
IV-41	-	VIII-46	649 249
IV-53	614 98	VIII-48	651 251
IV-62	-	VIII-66	665 262
IV-69	1924	VIII-74	-
IV-80	701 291	VIII-76	675 270
IV-82	196	X-04	-
IX-10	736 314	X-07	808 328
IX-12	-	X-15	814 329
IX-38	757 317	X-29	821 331
IX-39	758	X-34	-
IX-42	-	X-35	-
IX-48	764 319	X-54	837 334
IX-77	785 325	X-56	839 335
V-11	1188 404	X-68	1207 421
V-32	-	X-72	849 336
V-39	-	X-94	860 337
V-55	774 99	XI-07	-
V-80	726 311	XI-13	1209 423
V-94	-	XI-50	-
VI-07	931	XI-58	-
VI-34	-	XI-81	1212 426
VI-41	-	XII-07	1213 427
VI-48	891 355	XII-17	-
VI-49	-	XII-26	-
VI-52	-	XII-27	-
VI-55	897 359	XII-31	-
VI-65	-	XII-32	-
VI-70	1082	XII-35	1214 428

Clone ID	Sequence ID
XII-36	-
XII-52	-
XII-59	1216430
XIII-19	1219433
XIII-29	-
XIII-52	939378
XIII-62	-
XIII-84	-
XIII-92	1221435
XV-18	-
XV-22	1099388
XV-24	-
XV-25	1224436
XV-28	-
XV-34	-
XV-42	-
XV-68	-
XV-74	-
XV-93	-
XV-94	-
XV-96	-
XVI-36	1056382
XVI-53	1230439
XVI-59	-
XVI-66	1074384
XVI-76	1083386
XVI-77	1084387
XVII-07	-
XVII-08	-
XVII-17	-
XVII-28	-
XVII-29	-
XVII-31	1139392
XVII-36	-
XVII-39	-
XVII-40	1231440
XVII-48	1140393
XVII-55	-
XVII-58	-
XVII-67	-
XVII-72	-
XVII-76	1160394
XVII-82	-
XVII-87	1165395
XVII-95	1172396

Table 2b

List of sequences of probes informative for breast cancer

~~Please see the note at the bottom of Table 1. Some sequences are missing.~~

Clone ID	Sequence ID SEQ ID NO. in Sequence Listing
I-13	1331 444
I-14	1170 397
I-24	308 11
I-25	309 12
I-28	310 13
I-30	1180 398
I-37	1440 482
I-42	1332 445
I-48	321 19
I-54	1101 399
I-60	327 25
I-72	1335 446
I-81	338 31
I-82	339 32
I-86	1336 447
I-88	1182 400
I-95	1337 448
II-02	360 33
II-03	361 34
II-06	364 36
II-07	365 37
II-10	368 39
II-21	378 45
II-23	380 46
II-24	381 47
II-25	382 48
II-27	384 50
II-33	390 55
II-34	391 56
II-41	397 60
II-42	398 61
II-46	401 64
II-47	1338 449
II-48	403 66

II-52	406 68
II-57	411 73
II-58	412 74
II-59	413 75
II-60	414 76
II-61	415 77
II-62	416 78
II-64	418 80
II-67	421 83
II-69	423 85
II-70	424 86
II-74	428 90
II-80	434 96
II-82	436 98
II-84	438 99
II-87	441 100
II-88	442 101
II-96	450 105
III-01	452 106
III-02	453 107
III-06	458 109
III-08	460 111
III-12	463 114
III-13	464 115
III-17	4344 450
III-18	469 116
III-20	4183 401
III-21	471 117
III-23	473 119
III-24	474 120
III-25	475 121
III-26	476 122
III-27	477 123
III-28	478 124
III-29	479 125
III-32	482 127
III-33	483 128
III-35	485 130
III-39	487 131
III-40	488 132
III-42	489 133
III-45	492 135
III-46	493 136
III-47	494 137
III-48	495 138
III-56	503 144
III-57	504 145

III-58	505 <u>146</u>
III-59	506 <u>147</u>
III-61	507 <u>148</u>
III-62	508 <u>149</u>
III-63	509 <u>150</u>
III-64	510 <u>151</u>
III-66	512 <u>152</u>
III-67	513 <u>153</u>
III-70	515 <u>154</u>
III-74	518 <u>155</u>
III-75	519 <u>156</u>
III-78	521 <u>157</u>
III-80	523 <u>158</u>
III-81	524 <u>159</u>
III-82	1348 <u>451</u>
III-85	526 <u>161</u>
III-86	527 <u>162</u>
III-88	529 <u>163 + 164</u>
III-89	530 <u>165</u>
III-92	1351 <u>452</u>
III-93	532 <u>166</u>
III-95	534 <u>168</u>
III-96	1352 <u>452</u>
IV-04	682 <u>273</u>
IV-13	683 <u>274</u>
IV-14	684 <u>275</u>
IV-15	1185 <u>402</u>
IV-17	685 <u>276</u>
IV-23	1353 <u>454</u>
IV-26	1186 <u>403</u>
IV-31	687 <u>278</u>
IV-32	688 <u>279</u>
IV-35	1355 <u>455</u>
IV-37	664 <u>97</u>
IV-38	689 <u>280</u>
IV-42	691 <u>282</u>
IV-43	1239 <u>441</u>
IV-47	693 <u>284</u>
IV-53	614 <u>98</u>
IV-61	696 <u>286</u>
IV-64	697 <u>287</u>
IV-69	1924
IV-72	699 <u>289</u>
IV-80	701 <u>291</u>
IV-82	196
IV-85	702 <u>292</u>
IV-93	1360 <u>457</u>
IV-96	705 <u>295</u>

IX-10	<u>736314</u>
IX-12	738
IX-13	739 <u>315</u>
IX-24	747 <u>316</u>
IX-38	757 <u>317</u>
IX-39	758 <u>318</u>
IX-48	764 <u>319</u>
IX-50	766 <u>320</u>
IX-56	768 <u>321</u>
IX-62	773 <u>322</u>
IX-65	776 <u>323</u>
IX-72	782 <u>324</u>
IX-77	785 <u>325</u>
IX-91	796 <u>326</u>
IX-96	801 <u>327</u>
V-01	1361 <u>458</u>
V-03	706 <u>296</u>
V-04	707 <u>297</u>
V-07	708 <u>298</u>
V-08	709 <u>299</u>
V-11	1188 <u>404</u>
V-12	711 <u>301</u>
V-17	1364 <u>459</u>
V-24	714 <u>303</u>
V-25	1365 <u>460</u>
V-28	1189 <u>405</u>
V-38	1366 <u>461</u>
V-38	1190 <u>406</u>
V-39	1109 <u>389</u>
V-41	718 <u>305</u>
V-47	1368 <u>463</u>
V-49	1369 <u>464</u>
V-55	774 <u>99</u>
V-57	720 <u>307</u>
V-58	1370 <u>465</u>
V-61	721 <u>308</u>
V-64	722 <u>309</u>
V-65	1371 <u>466</u>
V-68	1448 <u>484</u>
V-71	1495 <u>496</u>
V-74	724 <u>310</u>
V-75	1372 <u>467</u>
V-80	726 <u>311</u>
V-90	1374 <u>468</u>
VI-03	864 <u>338</u>
VI-04	865 <u>339</u>
VI-07	931
VI-08	867 <u>340</u>

VI-09	1378 469
VI-12	869 341
VI-13	870 342
VI-14	871 343
VI-16	873 344
VI-19	875 345
VI-20	876 346
VI-21	1380 470
VI-23	878 347
VI-24	879 348
VI-25	1192 408
VI-26	881 349
VI-32	885 351
VI-39	887 352
VI-43	1382 471
VI-44	1193 409
VI-45	889 353
VI-48	891 355
VI-49	892 501
VI-50	893 356
VI-53	895 357
VI-55	897 359
VI-58	899 361
VI-66	903 363
VI-67	904 364
VI-70	1082
VI-71	1387 472
VI-74	905 365
VI-75	906 366
VI-76	907 367
VI-77	1103
VI-79	1389 473
VI-80	908 368
VI-85	910 369
VI-87	911 370
VI-88	912 371
VI-90	1390 474
VI-93	1391 475
VI-95	915 374
VI-96	1392 476
VII-02	1195 410
VII-03	1196 411
VII-06	1394 477
VII-08	1197 412
VII-09	1198 413
VII-10	1395 478
VII-11	1396 479
VII-15	1199 414

VII-17	560 <u>169</u>
VII-19	562 <u>171</u>
VII-21	564 <u>173</u>
VII-22	565 <u>174</u>
VII-23	566 <u>175</u>
VII-24	567 <u>176</u>
VII-25	1397 <u>480</u>
VII-26	250 <u>5</u>
VII-27	568 <u>177</u>
VII-29	570 <u>178</u>
VII-32	571 <u>179</u>
VII-33	575 <u>180</u>
VII-36	575 <u>182</u>
VII-39	576 <u>183</u>
VII-41	578 <u>185</u>
VII-42	579 <u>186</u>
VII-43	580 <u>187</u>
VII-46	583 <u>190</u>
VII-47	1200 <u>415</u>
VII-48	1201 <u>416</u>
VII-49	585 <u>191</u>
VII-54	589 <u>195</u>
VII-57	591 <u>197</u>
VII-58	592 <u>198</u>
VII-59	593 <u>199</u>
VII-62	594 <u>200</u>
VII-63	1202 <u>417</u>
VII-64	596 <u>202</u>
VII-66	598 <u>204</u>
VII-67	1399 <u>481</u>
VII-72	600 <u>206</u>
VII-73	601 <u>207</u>
VII-77	1203 <u>418</u>
VII-80	605 <u>210</u>
VII-82	607 <u>212</u>
VII-86	1453 <u>487</u>
VII-87	610 <u>214</u>
VII-90	612 <u>216</u>
VII-91	613 <u>217</u>
VII-92	614 <u>218</u>
VII-93	615 <u>219</u>
VII-96	617 <u>220</u>
VIII-09	618 <u>221</u>
VIII-10	619 <u>222</u>
VIII-13	622 <u>224</u>
VIII-16	624 <u>225</u>
VIII-20	628 <u>229</u>
VIII-21	629 <u>230</u>

VIII-22	1455
VIII-23	630 <u>231</u>
VIII-24	631 <u>232</u>
VIII-25	632 <u>233</u>
VIII-26	1456 <u>489</u>
VIII-27	633 <u>234</u>
VIII-28	634 <u>235</u>
VIII-29	635 <u>236</u>
VIII-30	636 <u>237</u>
VIII-31	637 <u>238</u>
VIII-32	638 <u>239</u>
VIII-33	639 <u>240</u>
VIII-34	1204 <u>419</u>
VIII-38	643 <u>243</u>
VIII-40	644 <u>244</u>
VIII-41	645 <u>245</u>
VIII-46	649 <u>249</u>
VIII-48	651 <u>251</u>
VIII-55	656 <u>256</u>
VIII-57	658 <u>258</u>
VIII-59	660 <u>259</u>
VIII-60	661 <u>260</u>
VIII-61	1205 <u>420</u>
VIII-64	663 <u>261</u>
VIII-66	665 <u>262</u>
VIII-73	672 <u>267</u>
VIII-74	673 <u>268</u>
VIII-76	675 <u>270</u>
VIII-80	679 <u>272</u>
X-07	808 <u>328</u>
X-15	814 <u>329</u>
X-20	817 <u>330</u>
X-29	821 <u>331</u>
X-34	825 <u>332</u>
X-46	833 <u>333</u>
X-54	837 <u>334</u>
X-56	839 <u>335</u>
X-68	1207 <u>421</u>
X-72	849 <u>336</u>
X-73	1208 <u>422</u>
X-94	860 <u>337</u>
XI-13	1209 <u>423</u>
XI-37	1460 <u>490</u>
XI-43	1210 <u>424</u>
XI-67	1211 <u>425</u>
XI-81	1212 <u>426</u>
XII-07	1213 <u>427</u>
XII-35	1214 <u>428</u>

XII-36	1215 <u>429</u>
XII-59	1216 <u>430</u>
XII-65	1028 <u>381</u>
XII-92	1217 <u>431</u>
XIII-03	917 <u>375</u>
XIII-04	1218 <u>432</u>
XIII-19	1219 <u>433</u>
XIII-24	926 <u>376</u>
XIII-51	938 <u>377</u>
XIII-52	939 <u>378</u>
XIII-67	947 <u>379</u>
XIII-69	949 <u>380</u>
XIII-88	1220 <u>434</u>
XIII-92	1221 <u>435</u>
XV-22	1099 <u>388</u>
XV-24	1101
XV-25	1222 <u>436</u>
XV-42	1108
XV-62	1226 <u>437</u>
XV-64	1118 <u>390</u>
XV-84	1125 <u>391</u>
XVI-19	1228 <u>438</u>
XVI-36	1056 <u>382</u>
XVI-53	1230 <u>439</u>
XVI-60	1071 <u>383</u>
XVI-66	1074 <u>384</u>
XVI-74	1081 <u>385</u>
XVI-76	1083 <u>386</u>
XVI-77	1084 <u>387</u>
XVII-31	1139 <u>392</u>
XVII-40	1231 <u>440</u>
XVII-48	1148 <u>393</u>
XVII-76	1160 <u>394</u>
XVII-87	1165 <u>395</u>
XVII-95	1172 <u>396</u>

Table 3

List of informative probes (Clone ID) selected for breast cancer diagnosis based on their occurrence criterion during variable selection

Occurrence*	Clone ID
100%	XI-8, XVI-66, VIII-66, XVI-59, VII-03, XIII-19, XII-35, X-35, XI-50, XII-26, IV-53, XIII-29, XIII-62, I-30, III-06, XV-22, XV-94, VII-15, VII-39, IX-39, XVII-39, III-40, VII-32
90%	I-52, VI-65, VI-34, IV-62, XV-34, XVII-58, V-11, VI-78, XII-36, XIII-92, VIII-29, XVI-53, XVI-77, XI-13, XIII-84, IV-14, XII-31, V-80, VII-48, XVII-29, XVII-72
80%	III-60, VIII-74, IX-12, X-04, XIII-52, VIII-30, IX-38
70%	VI-49, X-29, VIII-48
60%	IV-82, IX-10, VI-52, X-68, VII-77
50%	IV-15
40%	XV-28, II-70, V-55
30%	XVII-17, XVII-67
20%	XI-58, XVI-36, VIII-39, VIII-44, III-61, IV-69, XV-68, X-72
10%	IX-42, IX-77, X-94, XV-96, XVII-55
5%	XII-59, XVI-76, I-54, XV-18, V-94, X-54, VI-07, VII-47, XVII-31, XVII-87, XVII-48
In at least one model	II-41, VI-41, III-57, III-89, VII-73, XV-25, IV-26, X-34, IV-41, VII-90, XV-42, XVII-82, XII-27, VIII-20, I-28, VII-60, VIII-76, III-20, VI-84, XI-07, XVII-28, XII-17, XVII-36, XII-52, XVII-76, VIII-46, VI-70, XV-74, XV-93, VIII-31, II-87, V-39, VI-55, X-07, X-15, XII-07, XVII-07, XVII-08, XVII-95, I-24, IV-32, V-32, VI-48, VI-72, IV-80, IX-48, X-56, XV-24, XII-32, XVII-40

*100% = Genes appearing in all the 75 cross validated models; 90% = Additional genes appearing in at least 68 out of 75 cross validated models; 5% = Additional genes appearing in at least 4 out of 75 cross validated models and so on.

Table 4a

List of informative probes for diagnosis of Alzheimer disease

Clone ID	Sequence- ID SEQ ID NO. in Sequence Listing	Clone ID	Sequence- ID SEQ ID NO. in Sequence Listing
I-01	-	III-60	-
I-02	-	III-63	509 150
I-13	-	III-68	-
I-21	-	III-74	518 155
I-34	313 15	III-80	523 158
I-37	-	III-82	-
I-42	-	III-85	526 161
I-58	326 24	III-92	-
I-71	-	III-96	-
I-72	-	IV-23	-
I-86	-	IV-26	-
I-95	-	IV-29	-
II-03	361 34	IV-31	687 278
II-05	363 35	IV-34	-
II-06	364 36	IV-35	-
II-10	368 39	IV-45	-
II-24	381 47	IV-80	701 291
II-25	382 48	IV-82	-
II-26	383 49	IV-93	-
II-33	390 55	V-01	-
II-34	395 6	V-02	-
II-42	398 61	V-03	706 296
II-47	-	V-04	707 297
II-57	411 73	V-06	-
II-61	415 77	V-07	708 298
II-69	423 85	V-12	711 301
II-75	424 91	V-15	-
II-83	-	V-17	-
II-84	438 99	V-21	-
II-88	442 101	V-25	-
II-90	-	V-35	-
II-94	448 104	V-42	-
III-02	453 107	V-43	-
III-05	-	V-47	-
III-06	458 109	V-49	-
III-08	460 111	V-52	-
III-10	-	V-54	-
III-13	464 115	V-58	-
III-15	-	V-59	-
III-17	-	V-65	-
III-23	473 119	V-68	-
III-26	476 122	V-71	-
III-35	485 130	V-75	-
III-39	487 131	V-79	-
III-43	490 500	V-80	726 311
III-44	491 134	V-90	-
III-53	500 142	V-91	-
III-56	503 144	V-92	-

Clone ID	Sequence ID ID SEQ ID NO. in Sequence Listing
VI-02	-
VI-04	<u>865339</u>
VI-09	-
VI-10	-
VI-12	<u>869341</u>
VI-14	<u>871343</u>
VI-17	-
VI-20	<u>876346</u>
VI-21	-
VI-23	<u>878347</u>
VI-41	-
VI-42	-
VI-43	-
VI-44	-
VI-48	<u>891355</u>
VI-49	-
VI-50	<u>893356</u>
VI-53	<u>895357</u>
VI-71	-
VI-74	<u>905365</u>
VI-76	<u>907367</u>
VI-78	-
VI-79	-
VI-87	<u>911370</u>
VI-88	<u>912371</u>
VI-90	-
VI-93	-
VI-95	<u>915374</u>
VI-96	-
VII-02	-
VII-03	-
VII-06	-
VII-10	-
VII-11	-
VII-19	<u>563171</u>
VII-21	<u>564173</u>
VII-25	-
VII-36	<u>575182</u>
VII-42	<u>579186</u>
VII-43	<u>580187</u>
VII-46	<u>583190</u>
VII-59	<u>593199</u>
VII-63	<u>595201</u>
VII-66	<u>598204</u>
VII-67	-
VII-72	<u>600206</u>
VII-73	<u>601207</u>
VII-75	-
VI-02	-
VI-04	866
VI-09	-
VI-10	-
VI-12	<u>873344</u>
VI-14	<u>875345</u>
VI-17	-

[illegible]

Table 4b

List of sequences of probes informative for Alzheimer disease
~~Please see note to Table 1~~

Clone ID	Sequence ID SEQ ID NO. in Sequence Listing
I-09	298
I-10	299 <u>6</u>
I-15	300 <u>7</u>
I-16	301
I-17	302 <u>8</u>
I-19	304 <u>9</u>
I-20	305
I-22	306 <u>10</u>
I-23	307
I-24	308 <u>11</u>
I-25	309 <u>12</u>
I-28	310 <u>13</u>
I-31	311 <u>14</u>
I-32	312
I-34	313 <u>15</u>
I-38	314 <u>16</u>
I-39	315 <u>17</u>
I-40	316 <u>18</u>
I-44	317
I-45	318
I-46	319
I-47	320
I-48	321 <u>19</u>
I-49	322 <u>20</u>
I-53	323 <u>21</u>
I-56	324 <u>22</u>
I-57	325 <u>23</u>
I-58	326 <u>24</u>
I-60	327 <u>25</u>
I-64	328 <u>26</u>
I-67	330 <u>27</u>
I-69	331 <u>28</u>
I-71	332
I-72	333
I-73	334
I-77	335 <u>29</u>
I-79	336
I-80	337 <u>30</u>
I-81	338 <u>31</u>
I-82	339 <u>32</u>
VI-02	340
VI-03	341
VI-04	342

VI-06	343
VI-07	344
VI-08	345
VI-09	346
VI-11	347
VI-18	348
VI-19	349
VI-20	350
VI-21	351
VI-22	352
VI-25	353
VI-26	354
VI-27	355
VI-31	356
VI-33	357
VI-35	358
VI-48	359
II-02	360 <u>33</u>
II-03	361 <u>34</u>
II-05	363 <u>35</u>
II-06	364 <u>36</u>
II-07	365 <u>37</u>
II-08	366 <u>38</u>
II-09	367
II-10	368 <u>39</u>
II-11	369 <u>40</u>
II-12	370 <u>41</u>
II-13	371 <u>42</u>
II-14	372
II-15	373 <u>43</u>
II-16	374 <u>44</u>
II-17	375
II-18	376
II-20	377
II-21	378 <u>45</u>
II-22	379
II-23	380 <u>46</u>
II-24	381 <u>47</u>
II-25	382 <u>48</u>
II-26	383 <u>49</u>
II-27	384 <u>50</u>
II-28	385
II-29	386 <u>51</u>
II-30	387 <u>52</u>
II-31	388 <u>53</u>
II-32	389 <u>54</u>
II-33	390 <u>55</u>
II-34	391 <u>56</u>
II-35	392
II-37	393
II-38	394 <u>57</u>

II-39	395 <u>58</u>
II-40	396 <u>59</u>
II-41	397 <u>60</u>
II-42	398 <u>61</u>
II-43	399 <u>62</u>
II-44	400 <u>63</u>
II-46	401 <u>64</u>
II-47	402 <u>65</u>
II-48	403 <u>66</u>
II-49	404
II-50	405 <u>67</u>
II-52	406 <u>68</u>
II-53	407 <u>69</u>
II-54	408 <u>70</u>
II-55	409 <u>71</u>
II-56	410 <u>72</u>
II-57	411 <u>73</u>
II-58	412 <u>74</u>
II-59	413 <u>75</u>
II-60	414 <u>76</u>
II-61	415 <u>77</u>
II-62	416 <u>78</u>
II-63	417 <u>79</u>
II-64	418 <u>80</u>
II-65	419 <u>81</u>
II-66	420 <u>82</u>
II-67	421 <u>83</u>
II-68	422 <u>84</u>
II-69	423 <u>85</u>
II-70	424 <u>86</u>
II-71	425 <u>87</u>
II-72	426 <u>88</u>
II-73	427 <u>89</u>
II-74	428 <u>90</u>
II-75	429 <u>91</u>
II-76	430 <u>92</u>
II-77	431 <u>93</u>
II-78	432 <u>94</u>
II-79	433 <u>95</u>
II-80	434 <u>96</u>
II-81	435 <u>97</u>
II-82	436 <u>98</u>
II-83	437
II-84	438 <u>99</u>
II-85	439
II-86	440
II-87	441 <u>100</u>
II-88	442 <u>101</u>
II-89	443
II-90	444
II-91	445

II-92	446 <u>102</u>
II-93	447 <u>103</u>
II-94	448 <u>104</u>
II-95	449
II-96	450 <u>105</u>
III-01	452 <u>106</u>
III-02	453 <u>107</u>
III-03	454 <u>108</u>
III-04	455
III-05	457
III-06	458 <u>109</u>
III-07	459 <u>110</u>
III-08	460 <u>111</u>
III-09	461 <u>112</u>
III-11	462 <u>113</u>
III-12	463 <u>114</u>
III-13	464 <u>115</u>
III-14	465
III-15	466
III-16	467
III-17	468
III-18	469
III-19	470
III-21	471 <u>117</u>
III-22	472 <u>118</u>
III-23	473 <u>119</u>
III-24	474 <u>120</u>
III-25	475 <u>121</u>
III-26	476 <u>122</u>
III-27	477 <u>123</u>
III-28	478 <u>124</u>
III-29	479 <u>125</u>
III-31	481 <u>126</u>
III-32	482 <u>127</u>
III-33	483 <u>128</u>
III-34	484 <u>129</u>
III-35	485 <u>130</u>
III-37	486
III-39	487 <u>131</u>
III-40	488 <u>132</u>
III-42	489 <u>133</u>
III-43	490 <u>500</u>
III-44	491 <u>134</u>
III-45	492 <u>135</u>
III-46	493 <u>136</u>
III-47	494 <u>137</u>
III-48	495 <u>138</u>
III-49	496 <u>139</u>
III-50	497 <u>140</u>
III-51	498
III-52	499 <u>141</u>

III-53	500 <u>142</u>
III-54	501
III-55	502 <u>143</u>
III-56	503 <u>144</u>
III-57	504 <u>145</u>
III-58	505 <u>146</u>
III-59	506 <u>147</u>
III-61	507 <u>148</u>
III-62	508 <u>149</u>
III-63	509 <u>150</u>
III-64	510 <u>151</u>
III-65	511
III-66	512 <u>152</u>
III-67	513 <u>153</u>
III-69	514
III-70	515 <u>154</u>
III-71	516
III-73	517
III-74	518 <u>155</u>
III-75	519 <u>156</u>
III-77	520
III-78	521 <u>157</u>
III-79	522
III-80	523 <u>158</u>
III-81	524 <u>159</u>
III-83	525 <u>160</u>
III-85	526 <u>161</u>
III-86	527 <u>152</u>
III-87	528
III-88	529 <u>163/164</u>
III-89	530 <u>165</u>
III-91	531
III-93	532 <u>166</u>
III-94	533 <u>167</u>
III-95	534 <u>168</u>
III-96	535
VII-02	547
VII-03	548
VII-04	549
VII-05	550
VII-06	551
VII-07	552
VII-08	553
VII-09	554
VII-10	555
VII-11	556
VII-12	557
VII-14	558
VII-15	559
VII-17	560 <u>169</u>
VII-18	561 <u>170</u>

VII-19	562 <u>171</u>
VII-20	563 <u>172</u>
VII-21	564 <u>173</u>
VII-22	565 <u>174</u>
VII-23	566 <u>175</u>
VII-24	567 <u>176</u>
VII-27	568 <u>177</u>
VII-28	569
VII-29	570 <u>178</u>
VII-32	571 <u>179</u>
VII-33	572 <u>180</u>
VII-34	573
VII-35	574 <u>181</u>
VII-36	575 <u>182</u>
VII-39	576 <u>183</u>
VII-40	577 <u>184</u>
VII-41	578 <u>185</u>
VII-42	579 <u>186</u>
VII-43	580 <u>187</u>
VII-44	581 <u>188</u>
VII-45	582 <u>189</u>
VII-46	583 <u>190</u>
VII-48	584
VII-49	585 <u>191</u>
VII-50	586 <u>192</u>
VII-52	587 <u>193</u>
VII-53	588 <u>194</u>
VII-54	589 <u>195</u>
VII-55	590 <u>196</u>
VII-57	591 <u>197</u>
VII-58	592 <u>198</u>
VII-59	593 <u>199</u>
VII-62	594 <u>200</u>
VII-63	595 <u>201</u>
VII-64	596 <u>202</u>
VII-65	597 <u>203</u>
VII-66	598 <u>204</u>
VII-71	599 <u>205</u>
VII-72	600 <u>206</u>
VII-73	601 <u>207</u>
VII-74	602 <u>208</u>
VII-76	603 <u>209</u>
VII-77	604
VII-80	605 <u>210</u>
VII-81	606 <u>211</u>
VII-82	607 <u>212</u>
VII-83	608
VII-84	609 <u>213</u>
VII-87	610 <u>214</u>
VII-89	611 <u>215</u>
VII-90	612 <u>216</u>

VII-91	613 <u>217</u>
VII-92	614 <u>218</u>
VII-93	615 <u>219</u>
VII-94	616
VII-96	617 <u>220</u>
VIII-09	618 <u>221</u>
VII-10	619 <u>222</u>
VII-11	620
VII-12	621 <u>223</u>
VII-13	622 <u>224</u>
VII-15	623
VII-16	624 <u>225</u>
VII-17	625 <u>226</u>
VII-18	626 <u>227</u>
VII-19	627 <u>228</u>
VII-20	628 <u>229</u>
VII-21	629 <u>230</u>
VII-23	630 <u>231</u>
VII-24	631 <u>232</u>
VII-25	632 <u>233</u>
VII-28	634 <u>235</u>
VII-29	635 <u>236</u>
VII-30	636 <u>237</u>
VII-31	637 <u>238</u>
VII-32	638 <u>239</u>
VII-33	639 <u>240</u>
VII-34	640
VII-36	641 <u>241</u>
VII-37	642 <u>242</u>
VII-38	643 <u>243</u>
VII-40	644 <u>244</u>
VII-41	645 <u>245</u>
VII-42	646 <u>246</u>
VII-43	647 <u>247</u>
VII-45	648 <u>248</u>
VII-46	649 <u>249</u>
VII-47	650 <u>250</u>
VII-48	651 <u>251</u>
VII-50	652 <u>252</u>
VII-51	653 <u>253</u>
VII-53	654 <u>254</u>
VII-54	655 <u>255</u>
VII-55	656 <u>256</u>
VII-56	657 <u>257</u>
VII-57	658 <u>258</u>
VII-58	659
VII-59	660 <u>259</u>
VII-60	661 <u>260</u>
VII-61	662
VII-64	663 <u>261</u>
VII-65	664

VII-66	665 <u>262</u>
VII-67	666 <u>263</u>
VII-68	667
VII-69	668
VII-70	669 <u>264</u>
VII-71	670 <u>265</u>
VII-72	671 <u>266</u>
VII-73	672 <u>267</u>
VII-74	673 <u>268</u>
VII-75	674 <u>269</u>
VII-76	675 <u>270</u>
VII-77	676 <u>271</u>
VII-78	677
VII-79	678
VII-80	679 <u>272</u>
IV-02	681
IV-04	682 <u>273</u>
IV-13	683 <u>274</u>
IV-14	684 <u>275</u>
IV-17	685 <u>276</u>
IV-28	686 <u>277</u>
IV-31	687 <u>278</u>
IV-32	688 <u>279</u>
IV-38	689 <u>280</u>
IV-40	690 <u>281</u>
IV-42	691 <u>282</u>
IV-44	692 <u>283</u>
IV-47	693 <u>284</u>
IV-55	694 <u>285</u>
IV-56	695
IV-61	696 <u>286</u>
IV-64	697 <u>287</u>
IV-65	698 <u>288</u>
IV-72	699 <u>289</u>
IV-73	700 <u>290</u>
IV-80	701 <u>291</u>
IV-85	702 <u>292</u>
IV-93	703 <u>293</u>
IV-95	704 <u>294</u>
IV-96	705 <u>295</u>
V-03	706 <u>296</u>
V-04	707 <u>297</u>
V-07	708 <u>298</u>
V-08	709 <u>299</u>
V-09	710 <u>300</u>
V-12	711 <u>301</u>
V-18	712
V-20	713 <u>302</u>
V-24	714 <u>303</u>
V-37	716
V-40	717 <u>304</u>

V-41	718 <u>305</u>
V-48	719 <u>306</u>
V-57	720 <u>307</u>
V-61	721 <u>308</u>
V-64	722 <u>309</u>
V-65	723
V-74	724 <u>310</u>
V-80	726 <u>311</u>
V-81	727 <u>312</u>
V-87	728 <u>313</u>
VI-13	870 <u>342</u>
VI-14	871 <u>343</u>
VI-16	873 <u>344</u>
VI-23	878 <u>347</u>
VI-24	879 <u>348</u>
VI-28	883 <u>350</u>
VI-32	885 <u>351</u>
VI-38	886
VI-39	887 <u>352</u>
VI-45	889 <u>353</u>
VI-46	890 <u>354</u>
VI-49	892 <u>501</u>
VI-50	893 <u>356</u>
VI-52	894
VI-53	895 <u>357</u>
VI-54	896 <u>358</u>
VI-55	897 <u>359</u>
VI-57	898 <u>360</u>
VI-58	899 <u>361</u>
VI-63	900 <u>362</u>
VI-65	902
VI-66	903 <u>363</u>
VI-67	904 <u>364</u>
VI-74	905 <u>365</u>
VI-75	906 <u>366</u>
VI-76	907 <u>367</u>
VI-80	908 <u>368</u>
VI-81	909
VI-85	910 <u>369</u>
VI-87	911 <u>370</u>
VI-88	912 <u>371</u>
VI-91	913 <u>372</u>
VI-94	914 <u>373</u>
VI-95	915 <u>374</u>
VI-96	916
I-13	1177
I-14	1178 <u>397</u>
I-30	1180 <u>398</u>
I-54	1181 <u>399</u>
I-88	1182 <u>400</u>

III-20	1183 <u>401</u>
IV-15	1185 <u>402</u>
IV-26	1186 <u>403</u>
IV-62	1187
V-11	1188 <u>404</u>
IV-28	1189 <u>405</u>
IV-38	1190 <u>406</u>
IV-45	1191 <u>407</u>
VI-44	1192 <u>409</u>
VII-47	1200 <u>415</u>
I-42	1332 <u>445</u>
I-52	1333
I-86	1336 <u>447</u>
I-95	1337 <u>448</u>
III-10	1342
III-60	1347
III-82	1348 <u>451</u>
III-92	1351 <u>452</u>
IV-23	1353 <u>454</u>
IV-34	1354
IV-35	1355 <u>455</u>
IV-41	1356
IV-45	1357
IV-82	1359 <u>456</u>
V-01	1361 <u>458</u>
V-02	1362
V-06	1363
V-17	1364 <u>459</u>
V-25	1365 <u>460</u>
V-35	1366 <u>461</u>
V-42	1367 <u>462</u>
V-47	1368 <u>463</u>
V-49	1369 <u>464</u>
V-58	1370 <u>465</u>
V-75	1372 <u>467</u>
V-79	1373
V-90	1374 <u>468</u>
V-91	1375
V-94	1376
VI-10	1379
VI-41	1381
VI-43	1382 <u>471</u>
VI-71	1387 <u>472</u>
VI-72	1388
VI-79	1389 <u>473</u>
VI-90	1390 <u>474</u>
VI-93	1391 <u>475</u>
VII-25	1397 <u>480</u>
VII-60	1398
VII-67	1399 <u>481</u>
VIII-22	1403

VIII-26	1404
VIII-39	1405
VIII-44	1406
I-37	<u>1440482</u>
V-32	1445
V-52	<u>1447483</u>
V-68	<u>1448484</u>
V-92	<u>1449485</u>
VI-42	<u>1450486</u>
VI-78	1452
VII-86	<u>1453487</u>
VII-88	<u>1454488</u>
IV-29	<u>1490491</u>
V-15	<u>1491491</u>
V-39	<u>1492493</u>
V-54	<u>1493494</u>
V-59	<u>1494495</u>
V-71	<u>1495496</u>

Table 5

Samples

Diagnosis	No. of women
Normal/Benign	42*
DCIS	3
Invasive cancer	26

*From one woman, whole blood was collected at weeks 1,2,3,4,5 following menstruation. Hence, the number of unique normal/benign samples tested in the experiment is 75.

Information about women with breast cancer

Sample	AGE	Stage	Cancer type	Size hist. (mm)	Nodes
1	51	II	IDC	20	1/7
2	84	II	IDC	22	2/2
3	50	I	DCIS+ 1 IDC	>50 DCIS; 5 x 14	0/7
4	47	I	IDC	15	0
5	69	III	ILC g.2 + tubular adenocarcinoma	50 + 3	1 av 12 + 1 av 7
6	50	II	IDC	24	0
7	65	I	IDC	15	0
8	63	II	IDC	23	0
9	55	I	IDC + DCIS	4	0 av 1
10	52	0	DCIS + small colloid carcinoma foci	50 + 3	0
11	60	II	IDC	24	0
12	54	I	IDC	11	0
13		0	DCIS	20	0
14	49	0	DCIS	9	0
15	48	I	IDC	4	0
16	56	I	IDC	4	0
17	68	I	IDC	14	0
18	68	I	IDC	7	0
19	63	I	IDC	10	0
20	45	I	IDC	19	1
21	57	III	IDC	60	8/20

22	55	II	IDC/DCIS	35 + 55	0
23	71	I	IDC/extensive DCIS	8	0
24	56	I	IDC	9	?
25	66	II	IDC	26	0
26	66	I	IDC	15	?
27	61	I	IDC	9	?
28	?	?	?	?	?
29	65	I	IDC	11	0

Other diseases/conditions present in the women tested

Other diseases/conditions present in the women tested

Disease/condition
Diabetes
Asthma
Ulcerous colitis
Hemochromatose
Crohn's disease
Fibromyalgia
Psoriasis
Atopic eczema
Rheumatism
Allergies

Prior history of cancer in the women tested

Cancer type	No. of women
Breast	3
Colon	2
Stomach	1
Skin	1

Table 6

Number of samples tested by double cross validation and success of the diagnostic test
for breast cancer based on selected ionformative genes

Number of samples tested by double cross validation

Number of unique samples tested	75
Number of unique non cancer samples tested	46
Number of cancer samples tested	29

Success of the diagnostic test for breast cancer based on selected informative genes

Occurrence in percentage*	Number of informative probes	Specificity	Sensitivity	Accuracy	False Positive rate	False negative rate	Total error rate
100.00	23	84.78	75.86	81.33	15.22	24.14	18.67
90.00	44	91.30	79.31	86.67	8.70	20.69	13.33
80.00	51	86.96	79.31	84.00	13.04	20.69	16.00
70.00	54	89.13	75.86	84.00	10.87	24.14	16.00
60.00	58	89.13	75.86	84.00	10.87	24.14	16.00
50.00	59	89.13	75.86	84.00	10.87	24.14	16.00
40.00	63	89.13	75.86	84.00	10.87	24.14	16.00
30.00	66	86.96	75.86	82.67	13.04	24.14	17.33
20.00	74	89.13	75.86	84.00	10.87	24.14	16.00
10.00	79	89.13	75.86	84.00	10.87	24.14	16.00
5.00	90	86.96	79.31	84.00	13.04	20.69	16.00
1.33	139	84.78	72.41	80.00	15.22	27.59	20.00

*100% = Genes appearing in all the 75 cross validated models; 90% = Genes appearing in at least 68 out of 75 cross validated models; 5% = Genes appearing in at least 4 out of 75 cross validated models; and so on.

Table 7

Double cross-validation and details of the success of the diagnostic test for Alzheimer disease based on the expression 182 informative genes

Validation Result		Success of diagnostic test		
Total number of samples tested	14	Performance	Description	%
Number of Alzheimer's disease samples tested	7	Accuracy	Percentage of the total number of predictions that were correct	92.9
Number of Alzheimer's disease samples incorrectly predicted	1	Sensitivity	Percentage of positive cases that were correctly identified	85.7
Number of non-Alzheimer's disease samples tested	7	Specificity	Percentage of negatives cases that were correctly predicted	100
Number of non-Alzheimer's disease samples incorrectly predicted	0	False positive rate	Percentage of negatives cases that were incorrectly classified as positive	0.0
		False negative rate	Percentage of positive cases that were incorrectly classified as negative	14.3
		Total error rate	Percentage of the total cases incorrectly predicted	7.1

Table 8

Some relevant features of the blood donors. **B**, Female donors with breast cancer; **N**, Female donors with suspected mammogram but no breast cancer; **IDC**, invasive ductal carcinoma; **DCIS**, ductal carcinoma in situ; **na**, not available **nd**, not determined; **++**, no degradation of mRNA and no ribosomal contamination in the sample, **+**, no degradation of mRNA but ribosomal contamination in the sample.

		AGE	Cancer type/ breast abnormality	Size Hist. (mm)	mRNA Quality
1	B1	na	IDC	5	++
2	B2	49	DCIS	8	nd
3	B3	54	IDC	18	++
4	B4	59	IDC	12	+
5	B5	61	DCIS + micro invasive cancer	15+1.5	++
6	B6	55	IDC	12+17	nd
7	B6		IDC	12+17	nd
8	N1	45	Fibroadenoma	-	nd
9	N2	52	na	-	+
10	N3	55	Cyst	-	++
11	N4	54	na	-	++
12	N5	51	Benign ductal epithelium	-	nd
13	N6	57	Benign	-	nd

14	N7	50	na	-	++
15	N8	52	na	-	+

Table 9

List of sequences of probes informative for both alzheimer and breast cancer disease

Clone ID	Sequence ID ID NO. in Sequence Listing
I-24	308 11
I-25	309 12
I-28	310 13
I-48	321 19
I-60	327 25
I-72	333
I-81	338 31
I-82	339 32
II-02	360 33
II-03	361 34
II-06	364 36
II-07	365 37
II-10	368 39
II-21	678 45
II-23	680 46
II-24	381 47
II-25	382 48
II-27	384 50
II-33	390 55
II-34	391 56
II-41	397 60
II-42	398 61
II-46	401 64
II-47	402 65
II-48	403 66
II-52	406 68
II-57	411 73
II-58	412 74
II-59	413 75
II-60	414 76
II-61	415 77
II-62	416 78
II-64	418 80
II-67	421 83
II-69	423 85
II-70	424 86
II-74	428 90
II-80	434 96
II-82	436 98
II-84	438 99
II-87	441 100
II-88	442 101

Marked-Up Copy

II-96	450 <u>105</u>
III-01	452 <u>106</u>
III-02	453 <u>107</u>
III-06	458 <u>109</u>
III-08	460 <u>111</u>
III-12	463 <u>114</u>
III-13	464 <u>115</u>
III-17	468
III-18	469 <u>116</u>
III-21	471 <u>117</u>
III-23	473 <u>119</u>
III-24	474 <u>120</u>
III-25	475 <u>121</u>
III-26	476 <u>122</u>
III-27	477 <u>123</u>
III-28	478 <u>124</u>
III-29	479 <u>125</u>
III-32	482 <u>127</u>
III-33	483 <u>128</u>
III-35	485 <u>130</u>
III-39	487 <u>131</u>
III-40	488 <u>132</u>
III-42	489 <u>133</u>
III-45	492 <u>135</u>
III-46	493 <u>136</u>
III-47	494 <u>137</u>
III-48	495 <u>138</u>
III-56	503 <u>144</u>
III-57	504 <u>145</u>
III-58	505 <u>146</u>
III-59	506 <u>147</u>
III-61	507 <u>148</u>
III-62	508 <u>149</u>
III-63	509 <u>150</u>
III-64	510 <u>151</u>
III-66	512 <u>152</u>
III-67	513 <u>153</u>
III-70	515 <u>154</u>
III-74	518 <u>155</u>
III-5	519 <u>156</u>
III-78	521 <u>157</u>
III-80	523 <u>158</u>
III-81	524 <u>159</u>
III-85	526 <u>161</u>
III-86	527 <u>162</u>
III-88	529 <u>163/164</u>
III-89	530 <u>165</u>

Marked-Up Copy

III-93	<u>532166</u>
III-95	<u>534168</u>
III-96	535
IV-04	<u>682273</u>
IV-13	<u>683274</u>
IV-14	<u>684275</u>
IV-17	<u>685276</u>
IV-31	<u>687278</u>
IV-32	<u>688279</u>
IV-38	<u>689280</u>
IV-42	<u>691282</u>
IV-47	<u>693284</u>
IV-61	<u>696286</u>
IV-64	<u>697287</u>
IV-72	<u>699289</u>
IV-80	<u>701291</u>
IV-85	<u>702292</u>
IV-93	<u>703293</u>
IV-96	<u>705295</u>
V-03	<u>706296</u>
V-04	<u>707297</u>
V-07	<u>708298</u>
V-08	<u>709299</u>
V-12	<u>711301</u>
V-24	<u>714303</u>
V-41	<u>718305</u>
V-57	<u>720307</u>
V-61	<u>721308</u>
V-64	<u>722309</u>
V-65	723
V-74	<u>724310</u>
V-80	<u>726311</u>
VI-03	341
VI-04	342
VI-07	344
VI-08	345
VI-09	346
VI-12	<u>869341</u>
VI-14	<u>871343</u>
VI-19	349
VI-20	350
VI-21	351
VI-23	<u>878347</u>
VI-25	353
VI-26	354
VI-48	359
VI-50	<u>893356</u>

Marked-Up Copy

VI-53	895 <u>357</u>
VI-74	905 <u>365</u>
VI-76	907 <u>367</u>
VI-87	911 <u>370</u>
VI-88	912 <u>371</u>
VI-95	915 <u>374</u>
VII-02	547
VII-03	548
VII-06	551
VII-08	553
VII-08	554
VII-10	555
VII-11	556
VII-15	559
VII-17	560
VII-19	562 <u>171</u>
VII-21	564 <u>173</u>
VII-22	565 <u>174</u>
VII-23	566 <u>175</u>
VII-24	567 <u>176</u>
VII-27	568 <u>177</u>
VII-29	570 <u>178</u>
VII-32	571 <u>179</u>
VII-33	572 <u>180</u>
VII-36	575 <u>182</u>
VII-28	576 <u>183</u>
VII-41	578 <u>185</u>
VII-42	579 <u>186</u>
VII-43	580 <u>187</u>
VII-46	583 <u>190</u>
VII-48	584
VII-49	585 <u>191</u>
VII-54	589 <u>195</u>
VII-57	591 <u>197</u>
VII-58	592 <u>198</u>
VII-59	593 <u>199</u>
VII-62	594 <u>200</u>
VII-63	595 <u>201</u>
VII-64	596 <u>202</u>
VII-66	598 <u>204</u>
VII-72	600 <u>206</u>
VII-73	601 <u>207</u>
VII-77	604
VII-80	605 <u>210</u>
VII-82	607 <u>212</u>
VII-87	610 <u>214</u>
VII-90	612 <u>216</u>

Marked-Up Copy

VII-91	613 <u>217</u>
VII-92	614 <u>218</u>
VII-93	615 <u>219</u>
VII-96	617 <u>220</u>
VIII-09	618 <u>221</u>
VIII-10	619 <u>222</u>
VIII-13	622 <u>224</u>
VIII-16	624 <u>225</u>
VIII-20	628 <u>229</u>
VIII-21	629 <u>230</u>
VIII-23	630 <u>231</u>
VIII-24	631 <u>232</u>
VIII-25	632 <u>233</u>
VIII-28	634 <u>235</u>
VIII-29	635 <u>236</u>
VIII-30	636 <u>237</u>
VIII-31	637 <u>238</u>
VIII-32	638 <u>239</u>
VIII-33	639 <u>240</u>
VIII-34	640
VIII-38	643 <u>243</u>
VIII-40	644 <u>244</u>
VIII-41	645 <u>245</u>
VIII-46	649 <u>249</u>
VIII-48	651 <u>251</u>
VIII-55	656 <u>256</u>
VIII-57	658 <u>258</u>
VIII-59	660 <u>259</u>
VIII-60	661 <u>260</u>
VIII-61	662
VIII-64	663 <u>261</u>
VIII-66	665 <u>262</u>
VIII-73	672 <u>267</u>
VIII-74	673 <u>268</u>
VIII-76	675 <u>270</u>
VIII-80	679 <u>272</u>

Nucleotide sequences

~~Sequence ID -- 93~~SEQ ID NO: 1 nt: 405
GGATCCTGTGGCCACAGAGCTGCCCCAGCAGACGCTCCGCCCCACCCGGTGATGG
AGCCCCGGGGGACAATCGTGCCTGGGGAGGAGCAGGGTACAGCCCATTCCCCCAG
CCCTGGCTGACCTGGCCTAGCAGTTTGGCCCTGCTGGCCTTAGCAGGGAGACAGGG
GAGCAAAGAACGCCAAGCCGGAGGCCCGAGGCCAGCCGGCCTCTCGAGAGCCAGAG
CAGCAGTTGAATGTAATGCTGGGGACAGGCATGCTGCCGCCAGTAGGGCGGGGACC
CGGACAGCCAGGTGACTACCAGTCCTGGGGACACACTCACCATAAACACATCCCCA
GGCAGGACAGATCGGGGAAGGGGTGTGTACCAGGCTATGATTTCTCTTGCAATAAA
ATGTATTATTATT

~~Sequence ID -- 108~~SEQ ID NO: 2 nt: 550
GGCTTTGACAGAGTGCAAGACGATGACTTGCAAAATGTGCGCATCTGGAACGCAACA
TAGANACCATCATCAACACCTTCCACCAATACTCTGTGAAGCTGGGGCACCCAGAC
ACCCTGAACCAGGGGGAATTCAAAGAGCTGGTGCGAAAAGATCTGCAAAATTTTCT
CAAGAAGGAGAATAAGAATGAAAAGGTCATAGAACACATCATGGAGGACCTGGACA
CAAATGCAGACAAGCAGCTGAGCTTCGAGGAGTTCATCATGCTGATGGCGAGGCTA
ACCTGGGCCTCCCACGAGAAGATGCACGAGGGTGACGAGGGGCCCTGGCCACCACCA
TAAGCCAGGCCTCGGGGAGGGCACCCCCCTAAGACCACAGTGGCCAAGATCACAGTG
GCCACGGCCACGGCCACAGTCATGGTGGCCACGGCCACAGCCACTAATCAGGAGGC
CAGGCCACCCTGCCTNTACCCAACCAGGGCCCCCGGGCCTGTTATGTCAAACCTGTC
TTGGCTGTGGGGCTAGGGGCTGGGGCCAAATAAAGTCTCTTTCTCC

~~Sequence ID 110~~SEQ ID NO: 3
ACGAAGACAGACATCTGTGGAATGATTCACATCCTCTCAAGTTAGGAGGATGGAGG
CCTGCTTCATTAAGAAGCTGGGGGTAGGGTGGGGGTGGGGAGAACACTTAACAACA
TGGGGACCAGTCAGGGGAATCCCCTTATTTCTGTTTTGCATATGAGGAACCCTAGA
GCAGCCAGGTGAGGCTCTCTAGTTTAATAAAAATCATGGAAAGACTCTTAATGCAG
ACTCTTCTTAAGTGTTAATAGGGATTTTTTCAGCTTATTTTGGTTGCAGTTTCCAA
TTTTTAAAAATGTTGAGGTAATCTTTCCACCTTCCCAAACCTAATTCTTGTAGAT
GCATTAGTGTTGAACCAATGCTTTCTCATGTCTCAATTCTTTGTATATGCATTCTT
TTCAGATGTATTAAACAAACAAAAACCCTTC

~~Sequence ID -- 192~~SEQ ID NO: 4 nt: 286
CCGGTAATAGAAATAGAAAAGGGAGAGTGTCTTCATGCAATGTGGCATCCTGGATTG

Marked-Up Copy

GGTCTCGNNACAAAAACAGGACATTAGTGGGAAAATTGGAAATCTGAAAAAAGTCT
GAATTTTAGTTAATATAACCAATTTTCAGTCTCTTGGTTTTGACAGATGTACCATGGT
GATGTAAGATGTTGACCTTGGGGTAGGCTGGGTGAAGGGTATACAGGAACCTTTG
TACTATCTCTGCAACTTCTCTGTAAATCTAGTATCATTCCTCCAAAATAAAAGTTTATT
TAATTT

~~Sequence ID 250~~ SEQ ID NO: 5

GTGGAAGTGACATCGTCTTTAAACCCTGCGTGGCAATCCCTGACGCACCGCCGTGA
TGCCCAGGGAAGACAGGGCGACCTGGAAGTCCAACCTACTTCCTTAAGATCATCCAA
CTATTGGATGATTATCCGAAATGTTTCATTGTGGGAGCAGACAATGTGGGCTCCAA
GCAGATGCAGCAGATCCGCATGTCCCTTCGCGGGAAGGCTGTGGTGCTGATGGGCA
AGAACACCATGATGCGCAAGGCCATCCGAGGGCACCTGGAAAACAACCCAGCTCTG
GAGAACTGCTGCCTCATATCCGGGGGAATGTGGGCTTTGTGTTCACCAAGGAGGA
CCTCACTGAGATCAGGGACATGTTGCTGGCCAATAAGGTGCCAGCTGCTGCCCCGTG
CTGGTGCCATTGCCCCATGTGAAGTCACTGTGCCAGCCCAGAACACTGGTCTCGGG
CCCGAGAAGACCTCCTTTTTTCCAGGCTTTAGGTATCACCACTAAAATCTCCAGGGG
CACCATTGAAATCCTGAGTGATGTGCACTGATCAAGACTGG

~~Sequence ID 299~~ SEQ ID NO: 6

CAGCGCAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAAG
TTTTTTTCTCTTTGAAAGATAGAGATTGNTACAACCTACTTAAAAAATATAGTCAAT
AGGTTACTAAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAAG
ATTTTAAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAA
AGGTTTCTAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAATG
TATTTAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAGA
AGGGCAATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTT
AAAAGTTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAAGAGATTAAAC
CGAAGGTGATTAAAAGACCTTGAAATCCATGACGCANGGAGAATTGCGCATTTAAA
GCCTAGTTACGCATTTACTAAACGCAGACGAAAATGGGAAGATTAATTGGGAGTGG
TAGGATGAAACAATTTTGGAGAAGATAGAAG

~~Sequence ID 300~~ SEQ ID NO: 7

CTCAAAGGAGAAAAAAAACCTTGTAaaaaaaAGCAAAAATGACAACAGAAAAACAAT
CTTATTCCGAGCATTCCAGTAACTTTTTTGTGTATGTACTTAGCTGTACTATAAGT
AGTTGGTTTGTATGAGATGGTTAAAAAGGCCAAAGATAAAAGGTTTCTTTTTTTTT
CCTTTTTTGTCTATGAAGTTGCTGTTTATTTTTTTTTTGGCCTGTTTGATGTATGTGT

Marked-Up Copy

GAAACAATGTTGTCCAACAATAAACAGGAATTTTATTTTGCTGAGTTGTTCTAAAA
AAAAAAAAAAAAAAAAAAAA

~~Sequence ID 302~~ SEQ ID NO: 8

AGTAGAGACGGGGTTTCACTGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGT
GATCCGGCCACCTCGGCCTCCCGAAAGTGCTGGGATTACAGGCGTGAGCCACGGCG
CCCAGCCCCAGCCTGTCACTTAACTGATAAACGACAGATTAACAGTAGAAAAATT
TTATTTTGCATACATAATGAGGCTTCACAAAAGAGAAGTGAAAACCCAAGTAGGAG
TTTAGGGCTGGGGGCTTATATACCATTTAACAAGGGGTGATAAATTGTAAGAGAAT
AG

~~Sequence ID 304~~ SEQ ID NO: 9

TCCTTGGTTTCGATTTGTGGCAACAATCCAGTCTTTTTGTTTTTTTCAGGGATACC
ATATGTAACAGGTGCCATTGTTACTGTAACTTTTCACACATGCCTTCAGTTTGATG
TCAAAGTCATCATTTAGTGTAACAGCAAGTTATCTGTTAGGCTGCACATCATGAA
CTTTACTTTTAGAAAGTCTTATCTTTTATGCCACAGAAATAGCATTTGGCTATTAG
TCATGGATGGCAAAGAAATTAATTTTGAGTTGTTTGGATAAAAATGTTTCAGTTGA
CTGTAGTGTGTATTGAGAGACACTGCCAGTAAACAACTCTCTTGGTAGGTGGAAA
TCCCCTAGAAGTTACAGAAAATTGGGAGGAGGTGAACTTAATTAAATAACTTGAAT
TGTTTAGACATATTCAGAGCTTCTTATGACCTTGAAGAAATCACCCAACCTTCAAAA
GACCTCGGTTTTCTTCATTTGTAAAAATTAGGGAGTTTGACTAGATGTGTAAATCTAG
TTGTTAGTTAACTTCTAAGATGTAAAAACCCTCTTGTTTAAACAAAACCTACAAGA
TCAAGTTGCTTATCTGAAATCTTTATGAATCAACACTAGTCACTAAGTCTAGCTCG
ACC

~~Sequence ID 306~~ SEQ ID NO: 10

CTTTTCCTCCCGCTGTCCCCACGGAGGGGACTGCTCTCCCCGCTGCATCCTTTC
TGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAACCTC
GACATTTCGGGAAGTGTTTTTGAGAAGTCTCGGTTCGGTAAGGGAAGTCTTCCAAGTC
CGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTTCGGCCACCCCCCAGATGAGG
CAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTTCTCTCAGCTT
CCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGTCCATGCTGCCAT
CAAGGCATTTATTGCAGTGTAATTTTGCTTCCAAAGGATCAGGCCCTGAGAACAA
TGACCTTATTTCTACAACAGTGTCTGGGTTCGCTGCCAGCAGATGCCTCAGATAC
CAAGAGATAACAAAGCTGCAGCTCTTTTGATGCTGACCAAGAATGTGGATTTTGTG
AAGGATGCNCATGAANAAATGGACNAGCTGTG

Marked-Up Copy

~~Sequence ID - 308~~ SEQ ID NO: 11

nt: 373

AAGTGGGTCTTGCCATCCCTGAACTGNAATCATCCCTAACATATTCATACCTGTTT
TCATTTTAAAAGTTGGGTCAGTTTTTTTTATTAGTACATGTATTTCTATCCTACTGA
TTTATTTGCTATATCATCTAATTTAGTTTGAATATTCCATAATTTACTTAATTAGT
CCTGTATGGAGACCTAGCTCTTCTCAGTGTCTACTATTATAAACAATGCTACAGTG
AATATTGGTGNATAAATCCATACNCACCACGTACATATCTTAAGTTCTGGAAGAGA
TATTGCTAAACCAGAAGATAACCTGCATTTAAAATTTGACTGCTAGGGNCAGGGNC
ACATTTAATTAAATTAGAACAAANGAATGCATAATGNC

~~Sequence ID - 309~~ SEQ ID NO: 12

CCGGAATCGCGGCCGCGTCGACGAAAATATGTGCCCTGGCCAACTCCACAGGACTA
GTTCTAGGCAATCTGAAGGAAACCAGAAAATGTGAATTTCTCTTCCCTCAAAAAGC
TATACTGAAGTAGTATTTAATATTCAAGTACTTGTAATTTGCAGAACAGTACTTT
TTAATTTGACCCATGAATTCTATTTAAATTTGTCACTTAATATTTAGCCAAGAAGC
AAACCATCTAAAAAGATTTCTGGTTTATTTCTCCAACCTCCTAATAAATAGGGTCAC
ATATTTTTTAACTTTTTTCTAATTTGAAAAGTAATACAGGCATATGGTATTTTTAAA
AATGAAACAACACAAAGGGATATGTTTTGAAAAGTGGTCTTGCCATCCCTGAACTG
TAATCATCCCTAACATATTCATACCTGTTTTTCATTTTAAAAGTTGGGTCAGTTTTT
TTATTAGTACATGTATTTCTATCCTACTGATTTATTTGCTATATCATCTAATTTAG
TTTGAATATTCCATAATTTACTTAATTAGTCCTGTATGGAGACCTAGCTCTTCTCA
GTGTCTACTATTATAAACAATGCTACAGTGAATATTGGTGNATAAATCCTACACAC
CACGTAACATATCTTAAGTTCCTGGAAGAGATATTGCTAAACCAGAAGATAACCTG
CATTTAAAATTTGACTGCTAGGGTCAGGGTCACATTTAAATTAAATTAGAACAAAG
AATGCATAATGTCTTCGATAGCAATCTATTCAAGGTGCACCGTGGTCACAAAGGAA
AGCAAAACTGTC

~~Sequence ID - 310~~ SEQ ID NO: 13

nt: 564

CCTGGNCAGAGGCCTCTATCCTGTANTGATAATTGCCATCAAAATTGTCAAAAANG
ATTTAATTTCTATGGGNAATAGTCCTTTTCTTAGCTTCTGCCNNTCACTTGCTTAT
TTTTTGTGTGGGAATGGGGTTGGATAAACCAATGAACTTTATTATAAACAATCCC
ACCTATATCTANCAAATTTATATTTTCGGTGAAATACAGATATTTGCCTTTCTGGA
GTANTATAGAAGCTGTCAATATGTATCTACTGTACAGTACTAAATAGTATTCATTT
ATGAAATGAGTAGTGTTTGGGTGGCTGGGGTTAAGGAAAAATGAGACTTGGAATTG
TAGCTTTTATCCAAGTTTTGAGTATAAATAGGGTTTTGTTTTGTTTTTTTTAACCT
AAAAACTGAAATGCCATATAGAAAAACAGCATTGTTTTTACAGTTTGTAGTAAGTA
ACTTTTTAAAGATTTTATCAAAAAGAATTTTGTCTATNGTGAGTAAAAGAAGTTCT

Marked-Up Copy

AATAATGGCCTAATCACTGCATTTTTTAAAAACAAAGTTCAACACAAATGACATTT
GTTT

~~Sequence ID 311~~ SEQ ID NO: 14

CCTCTCCTCCATCTAAAGGCAACATTCCTTACCCATTAGTCTCAGAAATTGTCTTA
AGCAACAGCCCCAAATGCTGGCTGCCCCGGCCAAGCATTGGGGCCGCCATCCTGC
CTGGCACTGGCTGATGGGCACCTCTGTTGGTTCCATCAGCCAGAGCTCTGCCAAAG
GCCCCGCAGTCCCTCTCCCAGGAGGACCCTAGAGGCAATTAAATGATGTCCTGTTC
CATTGG

~~Sequence ID 313~~ SEQ ID NO: 15

nt: 554

CCCGGAATCGCGGCCCGCGTCGACAACAAACCTGCATGTTCTGCACATGTATCCAG
GAACTTAAAAAAAAAAAAAGATAGTTTGTGTGTCTTAATTGAATAATAGTAGATTT
ATAGATTAAAGATCTATGGGTTTTTAATATGGATTANAAATCTGTGGGTTTTTGAT
ATGGATTANAAATCTGTGGGTTTTTAATATGGATTGGAAATCTGTGGGTTTTTAAT
ATGGATTAAAAACATCTGTGGGTTTTTAATATGGATTAAACATCTGTGGGTTTTT
AATATGGATTAAACATCTGGGTTTTTAATATGGATTAAACATCTGTGGGTTTTTAA
TATGGGTTAAAAATCAAAAGAAAATGAACTATTTGCTCCAGTGCAGGAAAATACAG
GCAATACTGGATACAATTAGATGGTCAGGAGCGATAACCCGGTTGCCATTGTTTGA
AGAAGAGAATAAGGNGCTAGCATTCCCTATCCGTAGATAATTTGACAGCTAGGAAAT
AGGGGGAGTCTTCTATGTAGTTAGTGAAGGCTAAATGAACTATTATATGC

~~Sequence ID 314~~ SEQ ID NO: 16

CTTTTCCTCCCGCTGTCCCCACGGAGGGGACTGCTCTCCCCGCTGCATCCTTTC
TGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAACCTC
GACATTCGGGAAGTGTTTTTGAGAAGTCTCGGTCGGTAAGGGAAGTCTTCCAAGTC
CGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTTCGGCCACCCCCCAGATGAGG
CAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTTCTCTCAGCTT
CCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGTCCATGCTGCCAT
CAAGGCATTTATTGCAGTGACTATTTGCTTCCAAAGGATCAGGCCCTGAGAACAA
TGACCTTATTTTCTACAACAGTGTCTGGGTTGCGTGCCAGCAGATGCCTCAGATAC
CAAGAGATAACAAAGCTGCAGCTCTTTTGATGCTGACCAAGAATGTGGATTTTGTG
AAGGATGCACATGAAGAAATGGAGCAGGCTGTGGAAGAATGTGACCCTTACTCTGG
CCTCTTGAATGATACTGAGGAGAACAACCTCTGACAACCACAATCATGAGG

~~Sequence ID 315~~ SEQ ID NO: 17

Marked-Up Copy

TGGTACAGATACAAACTGGACTCTCAGGACAAAACGACACCAGCCAAACCAGCAGC
CCCTCAGCATCCAGCAGCATGAGCGGAGGCATTTTCCTTTTCTTCGTGGCCAATGC
CATAATCCACCTCTTCTGCTTCAGTTGAGGTGACACGTCTCAGCCTTAGCCCTGTG
CCCCCTGAAACAGCTGCCACCATCACTCGCAAGAGAATCCCCCTCCATCTTTGGGAG
GGGTGATGCCAGACATCACCAGGTGTAGAAAGTTGACAGGCAGTGCCATGGGGGC
AACAGCCAAAATAGGGGGGTAATGATGTACGGGCCAAGCACTGCCCAGCTGGGGGT
CAATAAAGTTACCCTTGTACTTG

~~Sequence ID 316~~SEQ ID NO: 18

CGCCACTTATCCAGTGAACCACTATCACGAAAAAACTCTACCTCTCTATACTAAT
CTCCCTACAAATCTCCTTAATTATAACATTCACAGCCACAGAACTAATCATATTAA
AA

~~Sequence ID 321~~SEQ ID NO: 19

CAGAACAGTACTTTTTAATTTGACCCATGAATTCTATTTAAATTTGTCACTTAATA
TTTAGCCAAGAAGCAAACCATCTAAAAAGATTTCTGGTTTATTTCTCCAACTCCTA
ATAAATAGGGTCACATATTTTTTAACTTTTTTCTAATTTGAAAAGTAATACAGGCA
TATGGTATTTTTAAAAATGAAACAACACAAAGGGATATGTTTTGAAAAGTGTTCTT
GCCATCCCTGAACTGTAATCATCCCTAACATATTCATACCTGTTTTCATTTTAAAA
GTTGGGTCAGTTTTTTTATTAGTACATGTATTTCTATCCTACTGATTTATTTGCTA
TATCATCTAATTTAGTTTGAATATTCCATAATTTACTTAATTAGTCCTGTATGGAG
ACCTAGCTCTTCTCAGTGTCTACTATTATAACAATGCTACAGTGAATATTGGTGT
ATAAATCCATACACACCACGTAACATATCTTAAGTTCCTGGAAGAGATATTGCTAA
ACCAGAAGATAACCTGCATTTAAATTTTGA CTGCTAGGGTCAGGGTCACATTTAA
ATTAAATTAGAACAAGGAATGCATAATGTCTTCGATAGCAATCTATTCCAGGTGCA
CCGTGGTCACAAAGGAAAGCAAACTGTCAATAACTTTCTTCTCA

~~Sequence ID 322~~SEQ ID NO: 20

TAGCATTTGGCCTTTTAAACATTTGTTTATTTTTTTTCTGAGAATGGCTAACACA
CTTTATTGAGGTTTCGAAATTAATAAAGAAAATAAAAGAAATGTATCTTCATTCATT
CTGTATGTTAGTGTTTTTAATTACCCCTTAGAATATATGGATAAAAAATACTATTCTT
TGTCTTGGAGAAGGTAAGAGTCTAGTTAGATGAATAAGGGTTATCTATGTAGAACA
ACTAGAGAATGAGAAGAGAGCTTATGAGATTGAGTACTACGTTATGCAGTAGAGTA
GCACGTCATCTGCTACTGAGTATGGTGTGATAACATTGTGTAACAGGAAAGTATGA
TCAATATCTACTTAAATTAAGGACAATATTAGCACTACATTGCTTTATTTTAAAG
TAAAAATTAGAGAACTAAACACAAGCATTGTAAGTACAATAAAAGCTGATCTTTCT

Marked-Up Copy

AGTTAAGCAGAATAATACATGTTCAAGCATCTGCTAAATCATTAATATAAGAATA
TAGGGGTTTTCTATAATCTTATTTTCTTTGGAAGAGTACCTCATTTTCAAGANGAG
AAGTTTCTAATTGCCACTTCTTTAAAAATAAAACAGGGTTTTAATGTTCCCAGCAC
AAAAATTAATATCTCTTCAAAAAGTCTCTTGTGATTAAAGTTGAATCCCTTGTCAT
ACTGCTTCTAATATTGACACTGACCTCCTTAGGTATTTTTCAGGGGTTATAATCTT
TTCTTAAGGTATCTTTTTTCAAGAATTGGATACCTTGGGCTT

~~Sequence ID 323~~SEQ ID NO: 21

CGCGTCGACTTTTAAAGTCATCTCTATAGGAAGGTGCTGGGCAGGGATCCCAGAGA
AAGAAAGGGTCCAAGACTCCATTAAGTCCCTGGATGAAGGGCACTGCTACAGCAG
CTAGTACCAGAGACTCTCCTATCTCACGGTTGAGGCAGACCCAGGATAGAATAGAG
AATAAAAGGAATGCTTATAGGAAACAATTTTGTATGGAATGCTAGATGGCCAAGCC
TCAGCCTTTGGTCCAGTGCAACCCTTGCCTCGCTTGTCAACAGTGAAAAATTAGTT
TGGTTAGAAGAACCATCTGGAAACACACCAGCTTCTGCTACCTTCATGCTCATTGT
TAAAAAAAGATTAACCAGTGTGAACATTCTGATCTGTTAATTCCAGGGACTGTTTT
CTTTCCAATGGACTGTTTGGTTGGTAGAATAACCCCCAAAAGCTCAAAGCTAAAATG
CATCATCAGTCCTAGTCGGCAGTTTCTTAAGAATGGACTGGCGGCGTGGTTGAGCT
GATATGGAAAAGCTGCACCTTCTGCAGAAGATCAACTGACCTGCTATCCCACCCC
AAATTC AACCTGAGGTATATTT CAGTGAAGCAGGTAGCTGTGCTTCTCAAAGCAGA
GAAGCAGTTTTTAAGAACC AAAAAGGTAGAGGAAATCTA

~~Sequence ID 324~~SEQ ID NO: 22

GTTTGTACAGGCAGAATTGGATAGATACAGCCCTACAAATGTATATGCCCTCCCC
TGAAAAAAATTGGATGAAAATCTGCACAGCAAAGTGAAACACACAGATAATAGGAA
CAAATGTAGTTCCCATGTGCCAAACAAAATAAATGAAATCTCTGCATGTTTGCAG
CATATCTGCCTTTTGGGAATGTAATCAAGGNATAATCTTTGGCTAGTGTTATGTGC
CTGTATTTTTTTTAAAATGGTACACCAGAAAAGGACTGGCAGTCTACTTCTACCATA
GTTAACTTCACCCTCTTTAATTT CACAACATATTCTTTGGAAGCAGGAAGAAATG
CTCATAAAGAGGATCAGACCTTCTTTCCCGTGAAACCAGTATTTGGCGCCATATAT
AAGCCTGGTTAAATTGGTCATCTAAAGCTGTCAAATAAGACATTCTGTGAAAGGTA
AACATCGAAACTGGTTATAAGTAAAACCATCAAGCCAACAACAGGGTCTTGAGATA
ACCTTTGAAGCTTATTGTCTGGCCTGCACCAGAAGATGTCTGCATTACTCATTGCT
AAAAATGTGTACACAGAACTGCACTAGGATTAATTGGTTCAAGAAGAAATTTAAAC
TTACGTTTGGGTTTCCATACAGCACTCTATTGAATACATGCATCTGAATTTAAGTT
GCAA

~~Sequence ID 325~~SEQ ID NO: 23

Marked-Up Copy

GACCAGTAATGGCTTTTAAGAGTCCATTTTGTTCATTGTCTCCCTAGTTAATTACAG
GTGGGGGATCTTTTGCCTCTATTCTCTTCATATTGAAATGAATCATACTCATGTTT
TGTGGAACCTCCTTAAAGTTGTAGCTGTCATGATCAGATTTTTTTTTATATTCCTCA
GCTTAACCTCTGCTACTTGATTTACAGTGACCCATAACCTACTCATCCTTGTTTAT
AGTGACACATAATCTTATCTCTTTATAGAACCTTAAATTTTATCATTATTTTCGCT
TAGAATACAGCATTTCTTTGCTTCTGTTGCTGGTTTGACTTAAGAAATAAGGCAGT
AACTCTGATCAATCAATTATCCATAAGGAAGGGCTTTTCATGGGTCTATTAATTT
GTTAGTACCCTAAGTATATCTGAAAAATATGTCTATTGAGAGAAGATTTTGGCATT
CCAGATGGTATAGTCTATATATATTTAAAGTTTTGAATTTGCTTATATATACTCAG
CTTTCTTTTTCTAGCATTTTTCATTTACCTGTTAATTGAAGTATACCCCCACAT
ATAAAAGTTCCTCTTAAAGACACTGGACTCTTTCTGGGGGGCTAAAATA

~~Sequence ID - 326~~SEQ ID NO: 24

nt: 554

CCCGGAATCGCGGCCCGCGTCGACAACAAACCTGCATGTTCTGCACATGTATCCAG
GAACTTAAAAAAAAAAAAAGATAGTTTGTGTGTCTTAATTGAATAATAGTAGATTT
ATAGATTAAAGATCTATGGGTTTTTAATATGGATTANAAATCTGTGGGTTTTTGAT
ATGGATTANAAATCTGTGGGTTTTTAATATGGATTGGAAATCTGTGGGTTTTTAAT
ATGGATTAAAAACATCTGTGGGTTTTTAATATGGATTAAACATCTGTGGGTTTTT
AATATGGATTAAACATCTGGGTTTTTAATATGGATTAAACATCTGTGGGTTTTTAA
TATGGGTTAAAAATCAAAAGAAAATGAACTATTTGCTCCAGTGCAGGAAAATACAG
GCAATACTGGATACAATTAGATGGTCAGGAGCGATAACCCGGTTGCCATTGTTTGA
AGAAGAGAATAAGNGCTAGCATTCCTATCCGTAGATAATTTGACAGCTAGGAAAT
AGGGGGAGTCTTCTATGTAGTTAGTGAAGGCTAAATGAACTATTATATGC

~~Sequence ID 327~~SEQ ID NO: 25

CGGCTACCGACAGAAGGACTATTTTCATCGCCACCCAGGGGCCACTGGCACACACGG
TTGAGGACTTCTGGAGGATGATCTGGGAGGGGAAGTCCCACACTATCGTGATGCTG
ACGGAGGTGCAGGAGAGAGAGCAGGATAAATGCTACCAAGTATTGGCCAACCGAGGG
CTCAGTTACTCATGGAGAAATAACGATTGAGATAAAGAATGATACCCTTTCAGAAG
CCATCAGTATACGAGACTTTCTGGTCACTCTCAATCAGCCCCAGGCCCGCCAGGAG
GAGCAGGTCCGAGTAGTGCGCCAGTTTCACTTCCACGGCTGGCCTGAGATCGGGAT
TCCCGCCGAGGGCAAAGGCATGATTGACCTCATCGCAGCCGTGCAGAAGCANCAGC
AGCAGACAGGCAACCACCCCATCACCGTGCAGTGCAGTGCCGGAGCTGGGCGAACA
GGTACATTCATAGCCCTCAGCAACATTTTGGAGCGAGTAAAAGCCGAGGGACTTTT
ANATGTATTTCAAGCTGTGAAGAGTTTACGACTTCAGAGACCACATATGGTGCAAC
CCTGGAACAGTATGAAATGTGCTACAAAGTGGTACAAGATTTATTGATATATTTCT

GATTATGCTAATTTCAATGAAGATCCTGCCTTAAATATTTTTTAATTTAATGGCAN
AT

~~Sequence ID 328~~SEQ ID NO: 26

CAAGACTCCATCTCAAAAAAAAAAAAAATCTACAGTGCTGAGTATATAAAATTAT
TAACACATTTTACAACAATATGTGTTTGTGGAGTTAAATATTTTTTGTCTTTAAAA
CAGGTAATTTTAGTGCATACTTAATTTGATGATTAAATATGGTAGAATTAAGCATT
TTAAATGTTAATGTTTGTTACATTGTTCAAGAAATAAGTAGAAATATATTCCTTTG
TTTTTTATTTAAATTTTTGTTCCTCTGTAAACTAAAAGAACACGAAGTAATTGGTC
ACAATTACTGGTGTTTAACTGCCAAATATGGGTAAATAAGGGAAAATTTTGTTTAA
TATTTAGTCCTTCTGAGATGGCTTGAATATTTGAATTTTGTGTACGTCTATACTG
GGTAGTCACAAGTCTTATAAACACTTTAGAGGAAAGATGGATTTTCAGTCTGTATTT
TTAAACATCATTTATTTTAAATCTGGTGCTGAAAAATAAGAAAAAAATTAACTGC
ATTCTGCTGTTCTTCTTTANAAGCATTCCCTGCGTAAATACTGCTGTAATACTGTCA
TGCAAAGTGTATCCTTTCTTGTCGTATCCTTTTTGGGGCAGTGGTTTTT

~~Sequence ID 330~~SEQ ID NO: 27

GCGGGAATCGCGGCCCGCGTCGACCTCAAAGGAGAAAAAAACCTTGTAAAAAAG
CAAAAATGACAACAGAAAAACAATCTTATTCCGAGCATTCAGTAACCTTTTTTGTG
TATGTACTTAGCTGTACTATAAGTAGTTGGTTTGTATGAGATGGTTAAAAAGGCCA
AAGATAAAAGGTTTCTTTTTTTTTTCTTTTTTGTCTATGAAGTTGCTGTTTATTTT
TTTTGGCCTGTTTGATGTATGTGTGAAACAATGTTGTCCAACAATAAACAGGAATT
TTATTTTGTCTGAGTTGTTCTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAATTTTAAATTTTTTAAATAAAACCCTTGTTTAT

~~Sequence ID 331~~SEQ ID NO: 28

GCCGCGTCGACCTGCATGAGCCACAGTTTCTTGACTGGAGGCCATCAACCCTCTTG
GTTGAGGCCTTGTTCTGAGCCCTGACATGTGCTTGGGCACTGGTGGGCCTGGGCTT
CTGAGGTGGCCTCCTGCCCTGATCAGGGACCCTCCCCGCTTTCCTGGGCCTCTCAG
TTGAACAAAGCAGCAAAACAAAGGCAGTTTTATATGAAAGATTANAAGCCTGGAAT
AATCAGGCTTTTTAAATGATGTAATCCCACCTGTAATAGCATAGGGATTTTGGAAG
CAGCTGCTGGTGGCTTGGGACATCANTGGGGCCAAGGGTTCTCTGTCCCTGGTTCA
ACTGTGATTTGGCTTTCCCGTGTCTTTCCTGGTGATGCCTTGTTTGGGGTTCTGTG
GGTTTGGGTGGGAAGAGGGCCATCTGCCTGAATGTAACTGCTAGCTCTCCGAAGC
CCTGCGGGCCTGGCTTGTGTGAGCGTGTGGACAGTGGTGGCCGCGCTGTGCCTGCT
CGTGTTGCCTACATGTCCCTGGCTTGTTGAGGCGCTGCTTCAACCTGCACCCCTCC

Marked-Up Copy

TTGTCTCATAGATGCTCCTTTTGACCTTTTCAAAATTAATATGGATGGGAAAGCTC
CTATGCCCTTTTGGCTTCCTGGTAGAAGGCGGGATGCCCAAGGGTCTGCCTGGGTGT
GGATTGGATGCTTGGGGTGTGGGGGTGGAAACTGTCTTGTGGCCCACTTGGGCCC
C

~~Sequence ID 335~~SEQ ID NO: 29

CCCGCGTCGACTTTTAAAGTCATCTCTATAGGAAGGTGCTGGGCAGGGATCCCAGA
GAAAGAAAGGGTCCAAGACTCCATTAAGTCCCTGGATGAAGGGCACTGCTACAGC
AGCTAGTACCAGAGACTCTCCTATCTCACGGTTGAGGCAGACCCAGGATAGAATAG
AGAATAAAAGGAATGCTTATAGGAAACAATTTTGTATGGAATGCTAGATGGCCAAG
CCTCAGCCTTTTGGTCCAGTGCAACCCTTGCCTCGCTTGTCAACAGTGAAAAATTAG
TTTGGTTAGAAGAACCATCTGGAAACACACCAGCTTCTGCTACCTTCATGCTCATT
GTTAAAAAAGATTAACCAAGTGTGAACATTCTGATCTGTTAATTCAGGGACTGTT
TTCTTTCCAATGGACTGTTTGTGGTAGAATAACCCCCAAAAGCTCAAAGCTAAAA
TGCATCATCAGTCCTAGTCGGCAGTTCCTTAAGAATGGACTGGCGGCGTGGGTGAG
CTGATTTGGAAAACCTGCCCTTCTGCAAAAAACACTGGCCTGCTTTCCA

~~Sequence ID 337~~SEQ ID NO: 30

CAAGACTCCATCTCAAAAAAAAAAAAAATCTACAGTGCTGAGTATATAAAATTAT
TAACACATTTTACAACAATATGTGTTTGTGGAGTTAAATATTTTTTGTCTTTAAAA
CAGGTAATTTTAGTGCACTTAATTTGATGATTAAATATGGTAGAATTAAGCATT
TTAAATGTTAATGTTTGTACATTGTTCAAGAAATAAGTAGAAATATATTCCTTTG
TTTTTTATTTAAATTTTTGTTCCCTCTGTAAACTAAAAGAACACGAAGTAATTGGTC
ACAATTACTGGTGTTTAACTGCCAAATATGGGTAAATAAGGGAAAATTTTGTTTAA
TATTTAGTCCTTCTGAGATGGCTTGAATATTTGAATTTTGTGTACGTCTATACTG
GGTAGTCACAAGTCTTATAAACACTTTAGAGGAAAGATGGATTTTCACTCTGTATTT
TTAAACATCATTTATTTTAAATCTGGTGCTGAAAAATAAGAAAAAATTAACTGC
ATTCTGCTGTTCTTCTTTAGAAGCATTCTGCGTAAATACTGCTGTAATACTGTCA
TGCAAAGTGTATCCTTTCTTGTCTATCCTTTTTTGGGGCAGTGGTT

~~Sequence ID 338~~SEQ ID NO: 31

CTGGACTGCATGACCAGATCTGATGGGTGAGACTCAGGTGGCATGGAAGAGCCGAA
AGAGGATACCATATGTGGGTGCCGGGGGGGATAGGTGAGAAGTACTAGAAGGCGGA
ATGGAAGGACACTTCTGCTCAGCTCTGTGACACGGGCAGGGACCCTGCAGGGCTCA
GGTCCTTTTAACACAGCAGCTTCATTCTAACACCAGCAGCGTTGGAACACACGTACA
AGTATGCAGACTAAGCTCTTGCTTGGCTGATACGGCTTTTTTGGGTTTTTAGAGAAC

Marked-Up Copy

ATGCATATATGTTCTCATTCATGGTACATGAACTCAGAAGCCTTACTGCCTATTTT
TGTTAATACTTCTGGGCAAACATTACCACTTACAACCTCACACCAGTTAGAAATCAT
TTGTAAAATGTTATTTAATAAAGCCAAAGAACTAAATCATATTTATTTTCCAAGGN
TTTCTAAGATCTCTGAACTAATGAGGTTTTTTTAAATCCCCATTAAGTACTCATCA
CTGCTAGTAAAAGCAGTTGTCTTTACCTTTAATTCCAGTGAGTCCCCTTAAATTTA
TTTTTTTATTATCTTTGGCTACATTGCCTTAGACAAAATGTGGTCACCCTAATTTAA
NGGATAAAATTACATCCTCACAGATTTCTTATTAAGAGGGTCTAANCCTTGAATA
ATCANCAGTGGAATGGAAGTCTTCTTTACTGGNTTTNATCCTTTCCCTTTTTTAT
CCCATG

~~Sequence ID 339~~ SEQ ID NO: 32

TTTTTTTTTAAATAAAGCTGTCGGCACTCAAGGGTAATTTTCATATCAGTGTGNTCT
ACAAGCTGGGGGAAAATGAGTTCTAATTGTTCANAGCTACCAAATCCTTCACCTTTA
GCATAAAGGTTTTAAAGATATCACAAAGATGCCAAGTGATTAATAATGTTTTAAACC
ACCCCTTTTTCTGTCTGAAAAACAATAAACAATATTACAACAGTATAGTTACA
GAAGGGTTCTATTTTCATATGTTTTATGCACACTGTGCCTCAAAGGTACTATTTAA
ATATATATACTTTTGAGGGGGTGGCTAATGCAGAAACACCCAAGACCTAAGGAAGA
TACAACCCCATTTCTAGGTGTGAGGTCTAAATGCTTCACACACCCACTTGTGACCT
TTTTTCATGAAGAATCATAACACTGTGCAGTGAGAAACAGTGGCAAAGCAATACTG
AAAGCATTTTAAATTATTTACTAGGTAAAAGGGTGAAGTACTTTAAATACAT
CAAATTTTCATCAT

~~Sequence ID 360~~ SEQ ID NO: 33

GCAAGTGAGAGCCGGACGGGCACTGGGCGACTCTGTGCCTCGCTGAGGAAAAATAA
CTAAACATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGC
ATTTTTTTGTGCAAACCTGTTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAG
TCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCT
AAAGAGAAAGGAAAATTTGAAGATATGGCAAAGCGGACAAGGCCCGTTATGAAAG
AGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAGAAGTTCAAGGATC
CCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGC
CCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAACT
GGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGG
CTGCGAAGCTGAAGGAAAAATACGAAAAGGTA

~~Sequence ID 361~~ SEQ ID NO: 34

nt: 622

CTGTNATNGAATCTGCTTGTNACTNAAATGCTAAACTCAATTCTGTAATTCAATAG

Marked-Up Copy

GTGCACCTNTCTGAGAAACATANNAGACAATGAGGAAAAGGATTCANCATTCCGTG
GAATTTGTACCATGATCAGTGTGAATCCCANTGGCGTAATCCAAGTAAGATGTTCA
CAAAGATTTGTTTTTAATGTCTAATTAATAAAATTTTAAAGGAAGAAACATTCTAA
TACTTTAATTATAAAAAGTTAACTATTTTCAAAGGTATCAAATACAGTTAAACCT
TTAAAATGTATATTTCTTAATATCTTGAAATTGTAATGCCTTTTTTTTTTTCCTAAA
TTTTTTTTTGTTCATGAAATGAGATAGTAACAGCAGATTGGGACAACAAGGTATATT
CTTGTCTTGAATCAGGCCATGGCTTCTTTCATCCAAATTTTCAGACCTCATTTATTT
ACTTTGTCCCTGCCTCCCATCCCTGGATATCANGTTTGTGGATATCTACAGTTAAT
AGAGTGACCAAATAGTAGGAATACTGTCTCTCTATTCTGAATAAAAATACTTTGAAT
CAGATTTAGAAATAATGAATAAAATACAAATCACCATTGAAATTGCTCTAATTTTG
AGAGCT

~~Sequence ID - 363~~SEQ ID NO: 35

nt: 628

ATCACNTGAGGCAAGAGTTTGAGCCAGCCTAGCTAACATGGTGAAACCCCATCTCT
ACAAAAATATAAAAATTAGCCTGGGTGGTGATGGGCACCTGTAACCCAGCTACTC
GGGAGGCTGAGGTAGGAGAATCACTTGAACCCGGGAGATGGAGGTTGCAGTGAGCC
AAGATCGTGCCACTGCACTCCAGCCTGTGTGACAGAACAAGACTCTGTCTCAAAAA
AAAATAATAATAATAATAATAATAAAAAGGAATAACATAGCTAGGAATAAATTTAA
TCAAAGAGGTGAAAGACTTATACACTTAAAACCTACAAAAAAAAAATCACTGAAGGA
ATTATAGACCCAAATAAAAATAAATAAAAAGACATTCTGTGTTTTAGGGAAAGAAG
ACTTAATATTGTTAAGATGTCAATACTACCCAAAGTGATCTACAGATTCAACATAA
TCCCTATCAAAATTCCAACAGCCTACTTTGTAGAAATGGAAAAGCCAATTTTCAAA
TTCAGATGGAATTGCGAGGGGTCTGAATAACAAAAACAATCTTGGGGAAAAAAA
CAAAAAACAAAGTCAAAGAACTCACACTTCTCTATTTATAAATTTACTACAAAGTT
ATAGTAATCAAA

~~Sequence ID - 364~~SEQ ID NO: 36

nt: 528

TGAACATCCAGCCATGTCATTTCTTCCATTCCCTGCCCTGGAGTAAAGTAGATTTAC
TGAGCTGATGACTTGTGTGCATTTGTACATTGCAACCTTAGCTTACCTCTTGAAGC
ATGTAGAGCATTCATCACCCACCATTCAATCACTGCCTACTCCCACCACAGCTGTT
TCGTGGTCTGTCTGCTCCCTGTGCCACCCCCACCCCATCAGGTGGGCCTTTTGCAA
GTGATGAAGTCACCTGTGGGGGAAGAGCTTTCCTTTCTCCTCAACTCAGAAGG
CCTCTTCTCTTGCTCAAGAGGGTGCTGCTGCTTTCTGCCTCCTTCCCCGGCCGGC
CTCCATCCCAGTTCACCTTTTCAGAAATGGCCCCTCAGTCAACTCTTCCCTTTTCT
CCTGGCTTTTTTATTTCTCCCAGTCTCTTAAGAGTATCCTTAGCTTTAAAAACAATA
ACACAGAGGATGGGTGCAGTGGCTCATGCCTGTAATCCCAGCACTTTGGAGCCTGG
GGCGGGCGGATCACTTGAGGNCA

~~Sequence ID 365~~SEQ ID NO: 37

GTCCCGGAATCGCGGCCGCGTCGACCTTTTCTATGCCTGCTATATAAACAGTACCT
TGCAAGATGTCCTGTCTGATATCCACAAAGGGGTATTGTCAACCCCAAGTTCAGAC
AGCTTTGTATTCTTCTGTCCCTGGATACATGAATTACTGCCATCTTTACACAGCGC
CCTAAATACCAACGCGAAGTTACCTGCTCAGCTTGAAGCTGCGCTGTACCCTGGA
ACCAGCACTTCTGCTGAATGACTCAGGATGAAGCCTCGACTTCTCCTTCCCATCCC
ATGCCAGACCCAGTGGCTCCTTTCCCAATCTGATCCAGTGAAGTCCAGC
TGTTGCAACCTGGGCATGAGGAGGAGTGCAAGATGGCTTTGTCCTACCTGGAAAGA
GGCTTTCTGGA

~~Sequence ID 366~~SEQ ID NO: 38

CACCATTTACACACAGTGGGTCCTTGAATAGCATCGTTTTATTCAATGTCATTTTG
TTATAACATTGAGAAAAAATTGATTCCCGGCTGGGGCCACTGTCTGTGCACCGT

~~Sequence ID 368~~SEQ ID NO: 39

nt: 329

GAAAGATCTAAAATCGACACCCTAACATCACAATTAAAAGAACTAGAGAAGCAAGA
GCAAATTCAAAGCTAGCAGAAGGCAAGAAATAACTAAGATCAGAGCAGAGCTGAA
AGAGATAGAGACACAAAAAACCATTCAAAAAAAACAATGAATCCAGGAGTTTTTT
TTTTAAAAGATCAACAGAATTGACAGACTGCTAGCAAGACTAATAAAGAAGAGAG
AAGCATCAAATAGACTCAATAAAAAATGATAAAGGGGATATCACCACCAATCCCAC
AGAAATACAACTACCATCAGAGAACACTATAAACACCTCTATGCAAAT

~~Sequence ID 369~~SEQ ID NO: 40

GAAAGATCTAAAATCGACACCCTAACATCACAATTAAAAGAACTAGAGAAGCAAGA
GCAAATTCAAAGCTAGCAGAAGGCAAGAAATAACTAAGATCAGAGCAGAGCTGAA
AGAGATAGAGACACAAAAAACCATTCAAAAAAAACAATGAATCCAGGAGTTTTTT
TTTTAAAAGATCAACAGAATTGACAGACTGCTAGCAAGACTAATAAAGAAGAGAG
AAGCATCAAATAGACTCAATAAAAAATGATAAAGGGGATATCACCACCAATCCCAC
AGAAATACAACTACCATCAGAGAACACTATAAACACCTCTATGCAAATAAAGTAG
AAAAT

~~Sequence ID 370~~SEQ ID NO: 41

GAAAGATCTAAAATCGACACCCTAACATCACAATTAAAAGAACTAGAGAAGCAAGA
GCAAATTCAAAGCTAGCAGAAGGCAAGAAATAACTAAGATCAGAGCAGAGCTGAA
AGAGATAGAGACACAAAAAACCATTCAAAAAAAACAATGAATCCAGGAGTTTTTT

TTTTAAAAAGATCAACA

~~Sequence ID 371~~ SEQ ID NO: 42

GCCCGGAATCGCGGCCGCTCGACGTAAGCTCGGCTGAATCCACGGTTCAAGAACA
GGAAAGAAGGCCAAGGCATAGGGAGTGGGGCAGTTGGGTGAATATTAGTACCTTTC
CCTCAGNTNCATTAATTACCCCTGCCTACTCTGCACAAAAGGATNTAACACAGTT
TCCTTTTTTAATGGCCAGGTACAGCTGCTTATATGGANGGGCATTNTNAATGATAT
CCTTNATCACTGTCTTAATCATCACATNCTTAAACAATCACTTTATTGTGTTAAG
GAAGATAAAAATGGCTGGGTTC AATTTCCGTTCTGGAAGAAATCGANTNAAAAGGT
AACCATTTAATAATGCANAGGGCANTTTCACTGCAGACCCTAATACTGGAAATTTT
TAAAAACAAATGAAAACTTCTACTTTTTCTTCTAAGCTTACTTAACCACCCAAAT
TTTCCAGCCACATATCTTCCTAGTCTACAACCTGCCTTTAACTTTAAGAGATGCTCA
AAAAAATGTAAATTCTCAAATACATTCTTATTACAATTACTGCTAACCT

~~Sequence ID 373~~ SEQ ID NO: 43

CCAGTGTGCTGGGATTACAGGCATGAGCCCTGCACCCAGCCTCTTAAACTGATCAT
ATGATATTGGTTCTCAACCAAGGGTGACTTTGCCCCCAGAGGATACTTGGCAATGT
CTGGAGATACTCAGTTGTCATGACTTGGACAGGTGCTACTGTCACCCAGTGGGTAG
AGGTCAGGGATGGTGCTAAACATAGGACAGCTGTCAAGAGAAAAGAATGTACCCAG
CCCCAAATGTCAGTAGGGCTGAGGTTGAGAAACCCAGCTGTAGCTGACGTGTGAAG
GACAGACTGGCCTGGAAGTGTGTTTTCTGCCCCCTTCCACCCCTGCATATTAGTTA
AGGCCAAAGGAAAAAAGGAATGCAGGAAATGCCCGTTAAAAATCTTCAAAACAATA
TAAATGATCAATTCCACTAAAACCCTTTACACATTTAAGTATAAAGGTATTGGTA
GGAAAATTTGTTATTCACTGCTTTTCTCAGTGTGTCATGAAATAATTATTTCTGCTGT
CAGTTT

~~Sequence ID 374~~ SEQ ID NO: 44

AAAAAAAAAATCACTGAAGGAATTATAGACCCAAATAAAAAATAAAAGACAT
TCTGTGTTTTAGGGAAAGAAGACTTAATATTGTTAAGATGTCAATACTACCCAAAG
TGATCTACAGATTCAACATAATCCCTATCAAATTTCCAACAGCCTACTTTGTAGAA
ATGGAAAAGCCAATTTTCAAATTCAGATGGAATTGCGAGGGGTTCTGAATAACAAA
AACAACTCTGGGGAAAAAAAACAAAAACAAAGTCAAAGA ACTCACACTTCTCTAT
TTATAATTTACTACAAAGTTATAGTAATCAAAGTCGACGCGGCCGCGATTCCGGG

~~Sequence ID 378~~ SEQ ID NO: 45

CGACTGCGGCTCTTCCTCGGGCAGCGGAAGCGGCGCGGGTCGGAGAAGTGGCCT
AAA ACTTCGGCGTTGGGTGAAAGAAAATGGCCCGAACCAAGCAGACTGCTCGTAAG

Marked-Up Copy

TCCACCGGTGGGAAAGCCCCCGCAAACAGCTGGCCACGAAAGCCGCCAGGAAAAG
CGCTCCCTCTACCGGCGGGGTGAAGAAGCCTCATCGCTACAGGCCCGGGACCGTGG
CGCTTCGAGAGATTTCGTCGTTATCAGAAGTCGACCGAGCTGCTCATCCGGAAGCTG
CCCTTCCAGAGGTTGGTGAGGGANATCGCCCAGG

~~Sequence ID 380~~SEQ ID NO: 46

GCAATTTAATTTTTAATAACAAAGATACTGTATTTTAACATGGTGAAATATACTTG
GCTAAGTCCAGATTAAAAAAGTATCTAGCCCAACAGTACAATTATACAGCT
TTGTACAGAACATTCCATAGATCAACAGAAAATACATTTGAGCGCAAAAATAAAAA
ATATTTAAGGAGAATCTCTAAGCAGCATTTTATTTCTGCAAAAGACATATCTTGTC
TGATTAAATATCTACAAGTGCTTTTCCTTTCAAAAATACATATATTCTTAATAGAC
TAAGTCATTAACAATGACCTGGTAATTCTTTCACCTTCAATTTGAATGATTTATAAG
CTAAATCTTCAACCACAAAAAGGTTTTTATTTGTATTAAGATGTTACCACTTTTGA
CAAAAAGCTTAAATATTTTATATTTCAAAGGAAAATTAGCAACATAACTTTACAA
TATATTCTATGATATTTTGATTGTGAGGGCTACTCTATTTAAACTGATGATCTCT
GTTGTGTTGCTCAGATGCAGGAAAGCAGCAAAA

~~Sequence ID 381~~SEQ ID NO: 47

nt: 534

GACTTANATCTAAATGGACCACATTCTCTACTTAAAAAATGCTATTAACCATGTG
ATCTTCTCAGTCATGAGGTAATCTGGTGACTACCCTTCCTCAAAGCCAGTTGGGAT
ATTCTTTGAATAGAGTAAACAGTGTTTCTAGGCTGGGAGACACCAGACATAGTTG
AGGACAGAGGTGCTAGAAAATAGGAAGTTTAAAAGCATGTGCGGTGATGCTCAGAG
GAGGTAAACCCCAACCTCATGCTCATAGCTTCCAATCATTTTCTCTAGTTCTTAAC
TCTTAAATGTGAGAAATGCTTGAAGATTCTAGTCATCTGAAGAAAGTCTCTTTATT
AAAGATTTTCATAAAAGAGACCAAAGCAGACAAACAGAAAAAGACATCTTGGGGAA
AAAAACAAGGATAATGGGAAGAGAAGGAAAGTTTTAAAAATTATCAATATCCTCAG
GGGGACAAAATATTATATCCTATAAAGACAGATTTTTATTTTTTAAAAAATAGAA
AGCAAAACAAGCTCCTAAAAATAAAGTTTG

~~Sequence ID 382~~SEQ ID NO: 48

nt: 444

GTTAAGGAAGTCAGCACTTACATTAAGAAAATTGGCTACAACCCCGACACAGTAGC
ATTTGTGCCAATTTCTGGTTGGAATGGTGACAACATGCTGGAGCCAAGTGCTAACA
TGCCTTGGTTCAAGGGATGGAAAGTCACCCGTAAGGATGGCAATGCCAGTGGAACC
ACGCTGCTTGAGGCTCTGGACTGCATCCTACCACCAACTCGTCCAAGTACCAAGCC
CTTGCGCCTGCCTCTCCAGGATGTCTACAAAATTGGTGGTATTGGTACTGTTCCCTG

Marked-Up Copy

TTGGCCGAGTGGAGACTGGTGTCTCAAACCCGGTATGGTGGTCACCTTTGCTCCA
GTCAACGTTACAACGGAAGTAAAATCTGTGAAATGCACCATGAAGCTTTGAGTGA
AGCTTTTCCTGGGGACAATGTGGGCTTCAATGTCAAGAATGTGTCTGTCAAG

~~Sequence ID - 383~~ SEQ ID NO: 49

nt: 566

CTTTGAAGAACTTTGCCAAATACTTTCTTACCAATCTCATGAGGAGAGGGAACATG
CTGAGAAACTGATGAAGCTGCAGAACCAACGAGGTGGCCGAATCTTCCTTCAGGAT
ATCAAGAAACCAGACTGTGATGACTGGGAGAGCGGGCTGAATGCAATGGAGTGTGC
ATTACATTTGGAAAAAATGTGAATCAGTCACTACTGGAAGTGCACAACTGGCCA
CTGACAAAAATGACCCCCATTTGTGTGACTTCATTGAGACACATTACCTGAATGAG
CAGGTGAAAGCCATCAAAGAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGG
AGCGCCCGAATCTGGCTTGGCGGAATATCTCTTTGACAAGCACACCCTGGGAGACA
GTGATAATGAAAGCTAAGCCTCGGGCTAATTTCCCATAGCCGTGGGGTGACTTCC
CTGGTCACCAAGGCAGTGCATGCATGTTGGGGTTTCCTTTACCTTTTCTATAAGTT
GTACCAAAACATCCACTTAAGTTCTTTGATTTGTCCATTCCTTCAAATAAAGAAAT
TTGGTA

~~Sequence ID 384~~ SEQ ID NO: 50

TTTTGGGGTTTATATATAAGCCTGGTTCTTGCTGAAACTGCTTATGTTGATAACCA
GTTAGTGAGTTCCTCTCTATTGACTTGCTGGGAAGTTTATAGAGACATTTTTTATG
CATTCAGAGATTTTCAGTACAAATCTTGAAAAAGGGACATTTAGGCCGGGCGCGGTG
GCTCACATCTGTAACCCTAGCACTCTGGGAGGCTGAGGTGGGTGGATCATGAAGTC
AAGAGATAGAGACCATCCTGGCAAAAATTAGCTGGGCGTGGTGGGGTGCGCCCGTA
GTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAGCCCGGGAGGCGGAG
GTTTCATTGAGCCGAGATAGTGCCACTGCACTCCAGCCTGGACAACAGAGCGAGAC
TGTGTCTT

~~Sequence ID 386~~ SEQ ID NO: 51

CTAAGGGTTTAAAGATGGAAAGAGGCATTGATGAACAGCTGGGGAAGGAGTAGTTT
GAGGTAGATGTGCAGATGGAATGAAGAGAAGGTCTCAAGAAGAGGGTGGAGCCAAA
GAGGGCTGCAGATTTAGAAGGCTAAAGTCTTTAGATGGCTTTGGATAGCCTGTTGT
ATCTTGACCATGCAGGTTACAGTGGAGCATGGAGTGGGGACAGAAGTGGAGGAAG
GAACCAGGGAACATGGAGTGAGAAGCTAAAGGAAAGTGATGCAGTAGATACATGGC
TCTAAAGTACTCAGGACTTTCAGAGGCTTAAACATAGGGTGACCAACTATCCCCT
ATGCCTGATACTAAGGGCATTCCCTGGATGTGGACCTTTCATTCCCCAAATTAGGA
AAGTCTTGGGCATACCAAGACAAGTTGGCCACCCTACTCAAAGTATGTAAGCTAA

Marked-Up Copy

CATATCTGTTCTCTAAGAGGTTAAAGCTGGATGGGGATACCAGATGTATGTACGTG
ATGCAGTTAAACAGCAATACAAGGGGGCAAGTCTACCTGATCGGCCAATTCAATGG
GA

~~Sequence ID 387~~SEQ ID NO: 52

GAAGCCAAACCAAAGGAGCTTCTACTTCATGATGCCATTTATGTAAAGTTCAGGCA
GAGAAAATCAGTGGTTTAAAGAAGTTAGAATAATGATTATCTTTGGAGGGATTGCAA
CTGGAAGAAGTCATGATTGGGATTTCTGGGTCCCTAATAGTGCTCTGTGTCTTGATC
TGAGTGCCGACTACATGAGTGGTTAGGTTTGCAAAATTCATTGAGTTATGCACTTA
ATGGTGTGTCTTATTAGAGCTGATGGAGGAGAGAGGGCTTCAATTTGCACAACTG
AGTAATCAGCTAGGCCCAGTCACTAGGTGAACAACCTTACTGCTCCAATCAGCCTTA
GAGCAGGAATCAAACCTCATGTCTCAGAAAAGTTATTAATTCAGCTTGTCTTGGGAC
TTCCTTCAGAGTCACTCTTGAATAGCTGAAATAGTAAATGTTAAATCTGTGGATGC
AAGTGTGTAAATTATTTTAGTCATCAGCTCTAATAAGATGGCCTTTGGGGAAATGA
GTATAAGGTCACGAAAATGAAATGGCAAGAAGGAGGTCTACTATTTCTTCTGTAAT
ACTGATTTTTTACCCCATCAGGGTCAGTCCCCAGAGGTTGTAAATGTGAAGCTTG-T
CTTTTTCTTTAATAA

~~Sequence ID 388~~SEQ ID NO: 53

CTTTGGACACTAGGAAAAAACCTTGTAGAGAGAGTAAAAAATTTAACACCCATAGT
AGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACACCCACTACCTA
AAAAATCCCAAACATATAACTGAACTCCTCACACCCAATTGGACCAATCTATCACC
CTATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCCTCCGCATAAG
CCTGCGTCAGATTAAACACTGAACTGACAATTAACAGCCCAATATCTACAATCAA
CCAACAAGTCATTATTACCCTCACTGTCAACCCAACACAGGCATGCTCATAAGGAA
AGGTTAAAAAAAGTAAAAGGAACTCGGCAAATCTTACCCCGCCTGTTTACCAAAAA
CATCACCTCTAGCATCACCAGTATTAGAGGCACCGCCTGCCCAGTGACACATGTTT
AACGGCCGCGGTACCCTAACCGTGCAAAGGTAGCATAATCACTTGTTCCCTTAATTA
GGGACCTGTATGAATGGCTCCACGAGGGTTCAGCTGTCTCTTACTTTTAACCAGTG
AAATTGACCTGCCCCGTGAAGAGGCGGGCATAACACAGCAAGACGAGAAGACCCTAT
GGAGCTTTAATTTATTAATGCAAACAGTCCTAACAAACCCCAGGTCCTAAACTCCA
AACCTGCATTAAA

~~Sequence ID 389~~SEQ ID NO: 54

CGACCCGGAATTCGCGGCCGCGTCGACTGAGTTCTTGACAAGAGTGTTTTTCCCTT
CCCGTCACAGAGTGGGCCCCACGACCTACGGCACTTTGACCCCGAGTTTACCGAAG

Marked-Up Copy

AGCCTGTCCCCAACTCCATTGGCAAGTCCCCTGACAGCGTCCTCGTCACAGCCAGC
GTCAAGGAAGCTGCCGAGGCTTTCCTAGGCTTTTCCTATGCGCCTCCACGGACTC
TTTCCTCTGAACCCTGTTAGGGCTTGGTTTTTAAAGGATTTTATGTGTGTTTCCGAA
TGTTTTAGTTAGCCTTTTGGTGGAGCCGCCAGCTGACAGGACATCTTACAAGAGAA
TTTGACATCTCTGGAAGCTTAGCAATCTTATTGCACACTGTTTCGCTGGAAGCTTT
TTGAAGAGCACATTCTCCTCAGTGAGCTCATGAGGTTTTTCATTTTTTATTCTTCCTT
CCAACGTGGTGCTATCTCTGAAACGAGCGTTAGAGTGCCGCCTTAGACGGAGGCAG
GAGTTTCGTTAGAAAGCGGACGCTGTTCT

~~Sequence ID - 390~~ SEQ ID NO: 55 nt: 523

GAATCCCTAGAAAAAGAGAATTCCCAACTTGATGAGGAAAACCTTAGAACTGCGAAG
GAATGTAGAATCTTTGAAGTGTGCAAGCATGAAAATGGCTCAGCTACAGCTAGAAA
ACAAAGAACTGGAAAGTGAAAAAGAGCAACTTAAGAAGGGTTTGGAGCTCCTGAAA
GCATCTTTCAAGAAAACAGAACGCTTAGAAGTTAGCTACCAGGGTTTAGATATAGA
AAATCAAAGACTGCAAAAACTTTAGAGAACAGCAATAAAAAATCCAGCAATTAG
AGAGTGAAC TACAAGACTTAGAGATGGAAAATCAAACATTGCAGAAAAACCTAGAA
GAACTAAAAATATCTAGCAAAAGACTAGAACAGCTGGAAAAAGAAAATAAATCATT
AGAGCAAGAGACTTCTCAACTGGAAAAGGATAAGAAACAATTGGAGAAGGAAAATA
AGAGACTCCGACANCAAGCAGAAATTAAAGATCCACATTTGAAGAAAATAATGTGA
AGATTGGAAATTTGGAAAA

~~Sequence ID - 391~~ SEQ ID NO: 56 nt: 566

CTTTGAAGAACTTTGCCAAATACTTTCTTACCAATCTCATGAGGAGAGGGAACATG
CTGAGAACTGATGAAGCTGCAGAACCAACGAGGTGGCCGAATCTTCCTTCAGGAT
ATCAAGAAACCAGACTGTGATGACTGGGAGAGCGGGCTGAATGCAATGGAGTGTGC
ATTACATTTGGAAAAAAATGTGAATCAGTCACTACTGGAAC TGCACAACTGGCCA
CTGACAAAAATGACCCCCATTTGTGTGACTTCATTGAGACACATTACCTGAATGAG
CAGGTGAAAGCCATCAAAGAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGG
AGCGCCCGAATCTGGCTTGGCGGAATATCTCTTTGACAAGCACACCCTGGGAGACA
GTGATAATGAAAGCTAAGCCTCGGGCTAATTTCCCCATAGCCGTGGGGTGA CTTC
CTGGTCACCAAGGCAGTG CATGCATGTTGGGGTTTTCTTTACCTTTTCTATAAGTT
GTACCAAAACATCCACTTAAGTTCTTTGATTTGTCCATTCTTCAAATAAAGAAAT
TTGGTA

~~Sequence ID - 394~~ SEQ ID NO: 57

GACCCGGAATCGCGGCCGCGTCGACCATTTTAGCCAAGGTGCCTCTATAGGGGTCA

Marked-Up Copy

AGACATCATGTGCCCAGACCTAAGGTCAGGAATGTCATATTTTTCTGTTAAAATCA
TTTTATTTCTGTGTATCTTACCTTTAAATCATTGTGGTTTACTCTGAGATTCTGTA
GTCCTAATATTGTATCATTGTGCTGTCTGCAAAACAACTTGAATCTATTTTGTTTG
CATCTTTTGTACATGTAACGCAGCTGTACTTTATGTTCTTTGCAACTGTTTCCAT
TATGAGAACGCTGTGCTATTTACAAGGTTACATTTTTCTTGGCCAGGCGAGGTGGT
CATGCCTGTGATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACTTGAGGTAA
AGAGTTGAGACCAGCCTGGCTAGCATGGCGAAGCCCAGTCTCTACTAAAAATACAA
AAATTGGCCGGGTGAAATTAGCCGGGCGTGGTGGTGTGTGCTTGTAAATCCCAGCTA
CTCGGGAGGCTGAGGCAGGAGAATCGCTTGAATCCGGGAGGCAGAGGTTGCAGTGA
GCCAAGATCANGCCACTGCACTCCACCTCGGGGTCAAGAGCGAAACTCTGTCTCAA

~~Sequence ID 395~~SEQ ID NO: 58

CCGTTTTAGTCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCTGCCTCGGCC
TCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCGTAAATCAGGTTT
TTTAAATGTTTGCCAAACCTTATCACTGACTTTTATAACAAAATTATTTACTATAA
TCATTAGGGAATATTTAAGTTCTGCTAATACTTAAAATTGCAGAGTGCTAAAACCA
GCAGTGAGTTTAGAATCAAGCTAAGCTTTATTGTTGCTACTATTTGAGGCATATTA
GTTGACTGGTGTTCATATGCAAGGCAGTCTACTGGGTGCAACAAGGGTTAGAAGGA
TATTTTTAAAAAACTGACCCTATTCTCAGGATGAAAATAATACACTAGTAATAGTC
TGCTCTGTTGGTTAACTCCTCGTAAGGAGGTCAATTAAAATGCTGTAGTGTTGCAA
GGGAAGGAGAGGAAGAATCATATTCCTTCACTAGCAGGATCAAGAAAGCTTTTATA
GAAATATACAAAATCTTCACTTCTTGAAGGATTGGTAAAATTTAATAGCCAACATT
GGGCACCTATTTCATTCTCTGAGTAAATATTTATTGCAT

~~Sequence ID 396~~SEQ ID NO: 59

CTTAAATCTAAATGGACCACATTCTCTACTTAAAAAAATGCTATTAACCATGTGAT
CTTCTCAGTCATGAGGTAATCTGGTGACTACCCTTCCTCAAAGCCAGTTGGGATAT
TCTTTGAATAGAGTAAAACAGTGTTTCTAGGCTGGGAGACACCAGACATAGTTGAG
GACAGAGGTGCTAGAAAATAGGAAGTTTAAAAGCATGTGCGGTGATGCTCAGAGGA
GGTAAACCCCAACCCTCATGCTCATAGCTTCCAATCATTTTCTCTAGTTCTTAACTC
TTAAATGTGAGAAATGCTTGAAGATTACTAGTCATCTGAAGAAAGTCTCTTTATTA
AAGATTTTCATAAAAGAGACCAAAGCAGACAAACAGAAAAAGACATCTTGGGGAAA
AAAACAAGGATAATGGGAAGAGAAGGAAAGTTTTAAAAATTATCAATATCCTCAGG
GGGACAAAATATTATATCCTATAAAGACAGATTTTTATTTTTTAAAAAATAGAAA
GCAAAACAAGCTCCTAAAAA

Marked-Up Copy

~~Sequence ID - 397~~ SEQ ID NO: 60

nt: 534

GACCCGGAATCGCGGCCGCGTCGACGGAAGCTCCTGCCCCCTCCTAAAGCTGAAGCC
AAAGCGAAGGCTTTAAAGGCCAAGAAGGCAGTGTTGAAAGGTGTCCACAGCCACAA
AAAGAAGGAGATCCGCACGTCACCCACCTTCCGGCGGCCGAAGACACTGCGACTCC
GGAGACAGCCCAAATATCCTCGGAAGAGCGCTCCCAGGAGAAACAAGCTTGACCAC
TATGCTATCATCAAGTTTCCGCTGACCACTGAGTCTGCCATGAAGAAGATAGAAGA
CAACAACACACTTGTGTTTATTGTGGATGTTAAAGCCAACAAGCACCAGATTAAAC
AGGCTGTGAAGAAGCTGTATGACATTGATGTGGCCAAGGTCAACACCCTGATTTCGG
CCTGATGGAGAGAAGAAGGCATATGTTTCGACTGGCTCCTGATTACGATGCTTTGGA
TGTTGCCAACAAAATTGGGATCATTTAAACTGAGTCCAGCTGCCTAATTCTGAATA
TATATATATATATATATCTTTTACCATAA

~~Sequence ID - 398~~ SEQ ID NO: 61

nt: 512

GGGAGCCCCCTCTTCCCTCAGTTGTTTCTACTCAGACTGTTGCACTCTAAACCTA
GGGAGGTTGAAGAATGAGACCCTTAGGTTTTTAACACGAATCCTGACACCACCATCT
ATAGGGTCCCAACTTGTTTATTGTAGGCAACCTTCCCTCTCTCCTTGGTGAAGAAC
ATCCCAAGCCAGAAAGAAGTTAACTACAGTGTTTTCTTTGCACCGATCCCCACCC
CAATTCAATCCCGGAAGGGACTTACTTAGGAAACCCTTCTTTACTAGATATCCTGG
CCCCCTGGGCTTGTGAACACCTCCTAGCCACATCACTACAGTACAGTGAGTGACCC
CAGCCTCCTGCCTACCCCAAGATGCCCCCTCCCCACCCTGACCGTGCTAACTGTGTG
TACATATATATTCTACATATATGTATATTAAACTGCACTGCCATGTCTGCCCTTT
TTTGTGGTGTCTAGCATTAACCTTATTGTCTAGGCCAAAGCGGGGGTGGGAGGGGAA
TGCCACAG

~~Sequence ID - 399~~ SEQ ID NO: 62

TTTTGGCATTACTTAATCCAATTATAAAAACTGAATTTTTTAAAAAACAGCACTTGT
TTTTTCTTCCAAGATTAATTTGAATTTTTTTATGGACATTAGAAAACATTGCAGTT
TAGTCATAATCAAAAATAAATCTTGAGGCTGGTAGAGCAGCTTTGTTGCTGTTTAT
ATTTTTATTGCTTACTGGATTTTCAAGTTTACCTAGTGCCATCAGTTTGGTATTTTG
CCACCTTGCACATTCAGTGATGTTTGAATTTTTCTTTTTCTTTTTTTCATATTACT
TTTAAATCCTGAATAGTTTGTGGCAGCTGGAGATCACCTAGTCCACCACTGTCCAA
CATGGCAATGGTAAGTAATATTGAGTAAAGAATAGAAAATTAGTAAAATGCATGGC
TTCAGAATTATAGCAATTTGCAAAATAGGTAAATGGATGAAAATTAGAATGACCAG
TTTAACTTTCCCCCAGCAGATTCTTCTGTAAACAATGCCCCTTCAAATAAAGG
AAGAACAAGTGGGTGTTATACCTATGTTATTTGGCTATGTTAGCACAATATGATGG
ACTAATTTGAGAAAAAGCATTTACTTCCTTTACTATTACTTCTTTTCTTTATAGGG

CTAAGTCTGCCTTCTGGGTCTTTGAA

~~Sequence ID 400~~ SEQ ID NO: 63

GAAGAAGCGCGAAGAGCCGTTAGTCATGCCGGTGTGGTGGCGGCGGCGGAGACTGC
GGGCCCCGTAGCTGGGCTCTGCGAGGTGCAAGAAAGCCTTTGAGGTGAAGGTGTATG
AAAGTCATCATAACAGATGTTTTCCAAAACTTGTAGAAGGTGTGAAAAAACTAC
TAGGATCACGCGGCATGTATTGAGCATATAGGTTGCTGTAGATGAATGTTCTTAGC
TGTCATGTTTAAAAATACTTCTGCTTCGTTACCTCAAGTGTGGCATGCAGCATTTT
GGAAGGAAAATTGAAGACGTGTTCAAGAAAACATGAACAGAAGCAAATGATGAAAA
TGAGCATTTTACTTGATGTTGATAACATCACAATAAATTATGGAGAAAAATACATA
TTTGGCTAACTTTTAATTGCTGAACAATAAAGTGTTTTCTTTTAAATCNAAAA

~~Sequence ID 401~~ SEQ ID NO: 64

GAAGCCAAACCAAAGGGAGCTTCTACTTCATGATGCCATTTATGTAAAGTTCAGGC
AGAGAAAATCAGTGGTTTAAGAAGTTAGAATAATGATTATCTTTGGAGGGATTGCA
ACTGGAAGAAGTCATGATTGGGATTTCTGGGTCCTAATAGTGCTCTGTGTCTTGAT
CTGAGTGCCGACTACATGAGTGGTTAGGTTTGCAAAATTCATTGAGTTATGCACTT
AATGGTGTGTCTTATTAGAGCTGATGGAGGAGAGAGGGCTTCAATTTGCACAAC
GAGTAATCAGCTAGGCCAGTCACTAGGTGAACAACCTACTGCTACCAATCAGCCT
TAGAGCAGGAATCAAACTCATGTCTCAGAAAAGTTATTAATTCAGCTTGTCTTGGG
ACTTCCTTCAGAGTCACTCTTGAATAGCTGAAATAGTAAATGTTAAATCTGTGGAT
GCAAGTGTGTAAATTATTTTAGTCATCAGCTCTAATAAGATGGCCTTTGGGGAAAT
GAGTATAAGGTCACGAAAATGAAATGGCAAGAAGGAGGTCTACTATTTCTTCTGTA
ATACTGATTTTTTACCCCATCAGGGTCAGTCCCCAAAGGTTGTAAATGTGAAGCTTG
GTCTTTTTCTTTA

~~Sequence ID 402~~ SEQ ID NO: 65

GACCCTATTCTCAGGATGAAAATAATACACTAGTAATAGTCTGCTCTGTTGGTTAA
CTCCTCGTAAGGAGGTACAATTAAAATGCTGTAGTGTTGCAAGGGAAGGAGAGGAA
GAATCATATTCCCTTCACTAGCAGGATCAAGAAAGCTTTTATAGAAATATACAAAAT
CTTCACTTCTTGAAGGATTGGTAAAATTTAATAGCCAACATTGGGCACCTTATTCAT
TCTCTGAGTAAATATTTATTGCATGCTTATCTTGTATCAACATTGNGATGAAAGCN
CAAGAATGAAAGAGGAGGGAGAATGTTTANAGAATAAGGCTGAAACACAGATTTTG
TAGGGAGCGTAGGGGAGACTGANAAAACAG

~~Sequence ID 403~~ SEQ ID NO: 66

Marked-Up Copy

AAGACACCTGATAGATTGTCTTGTATTATTTTTCCTTTGCCTTCTTACAATCTCAG
TGATTAGAATTGGGCTGAAAACAATACATCAAATTCTCAGCAAAATCCTTATGGGT
TGCTGGATAACCGAGGGTTTTTAAGATCTTTAGACTTCACTATATAGAACAAATGTT
GAATGGGAATTTTCTTTATTTCTATANCGTTTNG

~~Sequence ID 405~~SEQ ID NO: 67

CCCGGAATCGCGGCCGTCGACGATGAGCATTTTTTTCATGTGTCTTTTGGCTGCA
TAAATGTCTTCTTTTGAGAAGTGTGCGTTCATATCCTTTGCCCACTTTTTGATGGG
GTTGTTTTTTTTCTTGTAATTTGTTTGAGTTCATTGTAGATTCTGGATATTAGCCC
TTTGTCTAGATGAGTAGGTTGCGAAAATTTTCTCCCATTTTGTAGGTTGCCTGTTCA
CTCTGATGGTAGTTTCATTTGCTGTGCAGAAGCTCTTTAGTTTAATTAGATCCCAT
TTGTCAATTTTGGCTTTTGTGTCATTGCTTTTGGTGTTTTAGACTTGAAGTCCTT
GCCCATGCCTATGTCCTGAATGGTAATGCCTAGGTTTTCTTCTAGGGTTTTGATGG
TTTTAGGTCTAACGTTTCAGTCTTTAATCCATCTTTTAAAAGTCTCTTCACAGTAC
ATGAGTAGTAGTGACACCAATAATGTCAGAGCAGGGAACCTCCCAGGTTCTGCCCCAT
CCACAAAAACAACAAATAAGCTGGCAAAAACCTTAAGAATCAACTTTTGCAGATCT
CTGAAATCTAGTCAAACTTAAACAGAGGAAAGATTAATAAAGACNGGCTGCCTGA
GATAACACTAACACACAC

~~Sequence ID 406~~SEQ ID NO: 68

CATCAAATAAATAAATAAATAAATTTTAAAAGTCACAGCATTGAATTTTTTAAATGT
TTGGGATGATAAAGCACCTGCTTATCATGAAGCTANAGAAATTCAATGACACGTTT
GCCAGGGTCTTTGCTAGTGATGTTGGAACAAGTCTGTAATGCTGATGAAACATCAC
TGTTTCGGGCATTATTGCCCCAGAAAGACACTGACTGCAGCTGATGAAACAGCCCTT
CCAAGAATTAAGGATGCCAAAGACCAAATAACTGTGCTGAGATATACTTACGCAGC
AGGCATGCATAAGTGTAACCTTGCTGTTATAAGCAAAAGCTTGCGTTCTCACTGTT
TTCAAGGAGTGAATTTTCATACCAATCCATTATTATGCTAATAAAAAGGCATGGATC
ACCAGGGACATCTTTTCAGATTGGTTTTACAAACATTTTGTACCAGCAGCTTGTGC
TTACTGCAGGGAAGCTGACTGGATGATGACTGCAAGATTTTGTATATCTTAACAA
CTGTTGTGCTCATCCTCCAGCTGAAATTCTCATCAAAAATAATGTTTATGGCTCAC
ACCTGTAATCTCAACACTTTGGGAGGATTGCCTGACCCAGGAGTTCAAGCCCACCC
TGGGCAACACAGCAAGACCCAACCTNTC

~~Sequence ID 407~~SEQ ID NO: 69

TTTTAAAATCATAAAACGTTTCTTACAAAAGAGCATTACATTNTGCACACTGCTC
TGAACAGATGCCAGGGACATGTGGACTATTGTTACTTTTCCCTCCCTGTCCCACCCC

Marked-Up Copy

CCAAATGTTACAGTGACCACAAAGCAAGGTGTTACAATAATTACATGGGGGGAAT
TTTTTAAACCACCAACAATAACGAAAAATAAAATCCACTCACTCTGCTGCTGTTTC
AAAATTTCAATGTTAGTTTTTGCACGCCCTTCCCCCCCCCAACCCTGTTTGTAAGG
AACTAAAACATTACATCTGGTGAACAGCAAAGATTTCACTACACCTCAAATGCAGA
ACACCTATGAAGCAGAGGAATGTTGGCTTTTTTAAACAGAAGCAGATAAAAAAAAAA
GATGCAGGACTCCTTCAGTTCTTCACTAGTCTTAGAAAACTTTCCAGAATACTGC
TTCACACTATAAAAAAGAAAAAATATCTTGCATTAGAATCCTTCAACATCTGCATA
CTGCTTCACACTGTTTCGTTTCTAGGAGCACTTTGTACAGGACACTTCTGCTTATA
TTTCTTTAATCAGAACTTAGTTGGATGGGCCGGGCATGGTGGCTCACGCCTGTAAT
CCCAGCACTTTGGGAGGCCGAGG-GGGTGGATCACC

~~Sequence ID 408~~SEQ ID NO: 70

CCATCTCCAAATTTAGTATTCATTCTGTTTAGCATATTATCAGTTGCCATCTATTT
GTTTTAACTGATTACTTGAATCTGATTAAACATCACAGAAATGGGCTTTGATAAGA
ACAATATTGAATAAGAAATTTTAAATAACAAAACAGCTTATAGAAAAATTCAGCAT
AACTTTTCCATCACCTTCACCACCCCTTGCCTTTTATTATCCTGTCTGTATCACTG
CTTTCTGTTAGCAGTGTTGTGTGAGTTAGGATTTGGGCAGGAAAGCAAAAGCAACC
ACCCGTCATTTTCCCAGAATGAAGGGTTTGACGTAGGATGTAGACTTTGTATAGTA
GTTGGGAGAGCTGTGGGAGTGAAGGTCAGGGATGTCACCTACAGAAGTCAGGGAAT
CTGCCACCAGAGATCCTGCATCAGAAACAGCCAACAGCGTGCTTCTGAAGAACTAG
TGGGGAAGTGGCTATAATTCTTAGGAATCCCAGCAAGTCCGCACCACTGTCTCAGT
CTACAGCAGTGGAGAAAGGGGTTTCCAGGAGCTCTCTGGAAAGTTCCTGCCCACAC
TTTGCAACAATCTTCAGAGGATAATGGGCTTCTCTTCCAGCTTCCACACCCAACAA
GAGTGCTTTTCATCGGCCAACTCTAACCTGGAACCCTATGGCAGAGGGGATTTAGG
AGACAGTTTGTNATGTCTGTGGAATGCAAATGAANANGTANCAATGCTTANTTGAC
AGCGGNCATACACAAATNTNGAAA

~~Sequence ID 409~~SEQ ID NO: 71

GATCCGTNGACT

~~Sequence ID 410~~SEQ ID NO: 72

CTCTTCCCAGCCCCTGAGCCCAGCCCCTTCCCAAGTGGTGCCAGACAAAAAATAC
ATGGCCCTTTTCGTGTCTTGGGGGTGGAAAGGGAGGGATGAATTGGGGTGATAGAAC
CCTGGTGAATTCAGAGTAATCTTTCTTTAGAAAACTGGTGTTTTCTAAAGAAACAG
GATAGGAGTTTAGAGAAGGCACCAAAGCTTTCACTTTGGTTTGGCACCACTTTCTA
ACCATCTGTTTTTTCTACCCTAGCTATCTTTTATTGGTAAAATATAAATGTATAAT

Marked-Up Copy

TATGTTTGTAGAGCTTTACCAAGGAGTTTCCCTCCTTTTTTGTGTTGATTAGCA
AATTTTTGATTCTCCATTTTCCAAAAGTAAGAGACTCCAGCATGGCCTTCTGTTTG
CCCCGCAGTAAAGTAACTTCCATATAAAATGGTATTTGAAAGTGAGAGTTCATGAC
AACAGACCGTTTTCCATTTTCATCTGTATTTTATCTCCGTGACTCCACTTGTGGGTT
T

~~Sequence ID - 411~~SEQ ID NO: 73 nt: 505

TGGAGCTGAAAAATTCCTATTACCTAGGGGCATCACAACGCATTGCATTTGCCCCG
TGTTTGGGATGATGCTGGTGTAACCTACTATGCTGCCAGTCATGTAAAAGTATAG
CACACACAATTAGTAGGTAATGCTTGCAAATAATAATGAAAGACTCTGCTACTGGT
TTATGTATTTACTATGCTATACTTTTTTGTCTATTACTTTAGAGTGTACTCCTACTTT
TTTTTTTTTTTTTTTTTGTAGATGGAGTTTCACTCTTGTCCTGTAGGCTGGAGCGAAN
TGGCGCGATCTCGGCTTACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTG
CCTCANCTTCCCAGAGTAGCTGAGATTACAGGCATGCACCGCCACGCACGGGTAAT
TTTGTATTTTTTGGTAGAGACAGGGTTTACCATGTTGGCCAGGCTGGTCACCAACT
CCTGACCTCAGGTGACCCGCCTCCTCACCTCCAGAGTGTTGGGATTACAGGNGTGA
G

~~Sequence ID - 412~~SEQ ID NO: 74

ATAAAAAATTAGCTGGGGGTGATGGGGCCCTGTACCCCAGCTACTCGGGAGGTGAGGT
AGGAGAATCACTTGAACCCGGGAGATGGAGGTTGCAGTGAGCCAAGATCGTGCCAC
TGCACTCCAGCCTGTGTGACAGAACAAGACTCTGTCTCAAAAAAAAAATAATAATA
TAATAATAATAAAAAGGAATAACATAGCTAGGAATAAATTTAATCAAAGAGGTGAA
AGACTTATACACTTAAAACACAAAAAAAATCACTGAAGGAATTATAGACCCAA
ATAAAAAATAATAAAAAGACATTCTGTGTTTTAGGGAAAGAAGACTTAATATTGTT
AAGATGTCAATACTACCCAAAGTGATCTACAGATTCAACATAATCCCTATCAAAT
TCCAACAGCCTACTTTGTAGAAATGGAAAAGCCAATTTTCAAATTCAGATGGAATT
GCGAGGGGTTCTGAATAACAAAACACAATCTTGGGGAAAAAAAACAAAAACAAAG
TCAAAGAACTCACACTTCTCTATTTATAATTTACTACAAAGTTATAGNATCAAAGT
CGACGCGCCGCGATCCGGGC

~~Sequence ID - 413~~SEQ ID NO: 75

CACAGTACTCCATTTTGGGGTCCAACTGTAATGCTCAAAATAATAAATGCTTACA
CGAAAATTATTTATTGAGAATATTCATATAAAAATTACCTAAAGCAAAGTAAAAA
AGTAAATCAAGGTGGTATATTTGAAGTGAATGGTGATTGGAAATTTTGTAGCTGTA
ACAAAAAGAAAGAAACAACTTTTTTTAAAGCCTCATTCTCTTTTCTTTCAAATG

Marked-Up Copy

TACCTTATTCCCACACACTCTTGGGCTGACCTTTATTTTATCAATAAGCTCAATAT
TACTTTGTTTAAAATAAGATGCTTCAGCAAAAGTCATTCTCTCTTTAACCATATAA
TTTAAAACTCCTCTTCACGATTGATAGCAAAATCAGAAACGTTAGGGCACCAGTG
AGTTGAAAAAACTGGTCTTAAGTTGGAAAACTATTATTAATAATATTATCCTATC
CATCCATATCTATTGAAATTGTCAGGTCCATAATTTCATTTTAATTAATTATAGGA
AAGAAGAAAAGATAATACCCATTTGTTCTAT

~~Sequence ID 414~~SEQ ID NO: 76

CTCAGACTCTTTCTGCCCTAATGGCCATTACTATCCAGTCTGTATTGCTACAAGGG
ACCCACTGGTACCCCTTTTAGATTCTATCAAAAGGAACAGGGTTTTCTAGAGGCA
GGCAGCCTGGTGGTATGGCACAGCAGAAGCTTACTGCTAATGAAATGGGAACCTCC
CCCTCCCTTGTGGTTTCAGCACAGAACCTGAATGCCAGGAAAAATTCCTGGGCCAA
GAAGCTAAAGCTAAAGAAACCTTCCTTTTTTCAACGTTTTTTTTTCTTTCAAACCTG
TAGGGTCACTTTTGATTGAGGCAAAGGGTCTTACTGTAAGTGGAAAAGACTCACT
CCCCTAACATAAGTTTTCTACTGTGGTGGGATGGTGCCGCCGATATGCTTGATATG
CTTTTCCTTCCACATGTAAAGCTAGGAAACCTAACAGGATGTCAGCAGGGCAGTTA
ACTCTGGACTCANAGCCCTCAAGGGCATGTGGCANAACCTCATGGCATNCAAGACC
A

~~Sequence ID 415~~SEQ ID NO: 77

nt: 596

GTATAATTGATTCTTTTGAACCTAAAGTATAAGACTTCACGATTAGAAAAAATTA
TCCAAAGACTAATGTAATTAAGTGAGGAAAAGGTGCTGGAGGAACTGGATAACCAC
ATGGAAATGTATGAACCATGACCTCTATGTCACATACTATATATAAACTTAATTT
GAGGTGTATCACAGAGCTAACTGTGGGGGCTAAAACGTTGAAGCCTTTGGATGGCC
GCACAAGAGATGTCTGCATTCATAACCTTGGGGAGGGTATGAACATTTCTTGGTAA
CATGCAAAAAGCACTAACTGTAAAAGAGAACAGTTGGTCAGTTGAATTTTCATGAAA
CATTGTAACTTCTGCTAAACAACTGACACCATTAAGAATGTGGAAAAAGGCTGGG
CACAGTGGCTCATGCCTATAATCCCAGCATTTTGGGAGGCCGGGGCGGGAGAATCA
CTTGAGGCCAGGAGTTTGAAACCAGCCTGGGCAACATGGCAAGACCCCGACTCTAC
AAAAATATTTTTTAAAATTAGTTGGGTGTGGTGATGCACTCCTGTAGTCCTAGCTG
CCAGGANGCTAAGGNGGAAGGATCACTTAACCCTGG

~~Sequence ID 416~~SEQ ID NO: 78

CTGGTGGCGGCGGTTCGTGCGGACGCAAACATGCAGATCTTTGTGAAGACCCTCACT
GGCAAAACCATCACCTTGAGGTCGAGCCCAGTGACACCATTGAGAATGTCAAAGC
CAAAATTCAAGACAAGGAGGGTATCCACCTGACCAGCAGCGTCTGATATTTGCCG

Marked-Up Copy

GCAAACAGCTGGAGGATGGCCGCACTCTCTCAGACTACAACATCCAGAAAGAGTCC
ACCCTGCACCTGGTGTGCGCCTGCGAGGTGGCATTATTGAGCCTTCTCTCCGCCA
GCTTGCCCAGAAATACAACCTGCGACAAGATGATCTGCCGCAAGTGCTATGCTCGCC
TTCACCCCTCGTGCTGTCAACTGCCGCAAGAAGAAGTGTGGTCACACCAACAACCTG
CGTCCCAAGAAGAAGGTCAAATAAGGTTGTTCTTTCTTGAAGGGCAGCCTCCTGC
CCAGGCCCCGTGGCCCTGGAGCCTCAATAAAGTGTCCCTTTCATTGACTGGAGCAG

~~Sequence ID 417~~SEQ ID NO: 79

GCAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCCGCAGATAAGTTT
TTTTCTCTTTGAAAGATAGAGATTAATACAACCTACTTAAAAATATAGTCAATAGG
TTACTAAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAAGATT
TTAAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAAAGG
TTTCTAAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAATGTAT
TTAAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAGAAGG
GCAATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTTAAA
AGTTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAAGAGATTAAACCGA
AGGTGATTAAAAGACCTTGAAATCCATGACGCAGGGAGAATTGCGTCATTTAAAGC
CTAGTTAACGCATTTACTAAACGCAGACCAAAATGGAAAGATTAATTGGGAGTGGT
AGGA

~~Sequence ID 418~~SEQ ID NO: 80

CCCGGAATCGCGGCCGCGTCGACGGGAGGTGATAGCATTGCTTTCGTGTAAATTAT
GTAATGCAAAATTTTTTTAATCTTCGCCTTAATACTTTTTTATTTTGTTTTATTTT
GAATGATGAGCCTTCGTGCCCCCCTTCCCCCTTTTTTGTCCCCCACTTGAGATG
TATGAAGGCTTTTGGTCTCCCTGGGAGTGGGTGGAGGCAGCCAGGGCTTACCTGTA
CACTGACTTGAGACCAGTTGAATAAAAGTGCACACCTTATAAAAAA

~~Sequence ID 419~~SEQ ID NO: 81

CCCGGAATCGCGGCCGCGTCGACGGGAGGTGATAGCATTGCTTTCGTGTAAATTAT
GTAATGCAAAATTTTTTTAATCTTCGCCTTAATACTTTTTTATTTTGTTTTATTTT
GAATGATGAGCCTTCGTGCCCCCCTTCCCCCTTTTTTGTCCCCCACTTGAGATG
TATGAAGGCTTTTGGTCTCCCTGGGAGTGGGTGGAGGCAGCCAGGGCTTACCTGTA
CACTGACTTGAGACCAGTTGAATAAAAGTGCACACCTTATAAAAA

~~Sequence ID 420~~SEQ ID NO: 82

CTTCATTTGAAATGGTTGAATCTGCTGTGTAATAAAGTGGTTCAACCATGATTAGG

Marked-Up Copy

AACTGAAATTTAGTAGAAGAGGGAAAAGGAGTTAATGTAACAAATTATTTTAGCTA
CAAACCCCGGTAATAGAGCACTTGGGGGATGGGATGGGGTGGGTTGGTGAGACAAT
CAGAATGGTAAATTGATTAAATGCTCCTAACCTGTAATTTTGTGCATAGAGCACC
CTATGCTGTGGAAATAACTGTTCTTAGATTTTCATTGTAAGTGGACTGTTTCAGGTTG
CCCAGAGGGAAAGAACATTCTAATTCTAATAAAATAAACTTTTATTTTGTTTA

~~Sequence ID 421~~SEQ ID NO: 83

TGTCATTGAATCTGCTTGTTACTTAAATGCTAAACTCAATTCTGTAATTCAATAGG
TGCACCTCTCTGAGAAACATAAGAGACAATGAGGAAAAGGATTCAGCATTCCGTGG
AATTTGTACCATGATCAGTGTGAATCCCAGTGGCGTAATCCAAGTAAGATGTTTAC
AAAGATTTGTTTTTAATGTCTAATTAATAAAATTTTAAAGGAAGAAACATTCTAAT
ACTTTAATTATAAAAAGTTAACTATTTTCAAAGGTATCAAATACAGTTAAACCTT
TAAAATGTATATTTCTTAATATCTTGAAATTGTAATGCCTTTTTTTTTTCTCTAAAT
TTTTTTTGTGCATGAAATGAGATAGTAACAGCAGATTGGGACAACAAGGTTATATTC
TTGTCTTGAATCAGGCCATGGCTTCTTTCATCCAAATTTTCAGACCTCATTTATTTA
CTTTGTCCCTGCCTCCCATCCCTGGATATCAGTTTGTGGATATCTACAGTTAATAG
AGTGACCAAATAGTAGGAATACTGTCTCTCTATTCTGAATAAAATCTTTGAATCAG
ATTTAGAAATAATGAATAAAATACAAATCAGCCATTGAAATTGCTCTAATTTTGAG
AGCTTATGATTTATTCATCTTTGGTTTCCAAGTTCAAGTTATATGTAGACATTTTA
ATT

~~Sequence ID 422~~SEQ ID NO: 84

GCTTCCTAGGTGAGGTCACGAGGAAACCTGCTGGCCAAGTGACCTGGCAGGGTGTG
GCCAGTGTGGCCAGGGCCGCCGAGCCTGCTTTCCTTCCCTGCAGCAGGAACCCTTC
TGGGGCTGTGATCCTGCGATGGTGCCTGGGTGGGAGTGGGGGTGGGGGGCGGGATG
GTCTCCCTACCTGCCAGCTTCTTGTTTTGAGGTGAGGACAGCCCCGGAAGCTCANA
CTTGGCTCCTGTCCATGTACTTGGGGCCATGAGCTCTGCAGGGACCTTGGAAGAN
AGAGACGGGTGGTGTANGGCANGGGAAGGCATTGTCTTCAAACAGGAAAAAGCTGA
NAATGGAAACAGGCGAAACTTACCAAGTGTAACATCACCTGGAAGTGAAGGAGGGT
GGGAAGGTTTTAATTATTTTAAAAATAGAGATGGGGTCTCACTATGTTGCCAGGC
TGGTCTCAAACACTACTGGGCTCAAGTGAACCTCCTTCT

~~Sequence ID 423~~SEQ ID NO: 85

nt: 387

TGTTTCTCNAGGGCGAGAGGCTGTCTTANAGCACCATTCTCTGGCCCTNGTCCCAT
GAGAAGGAACCGCACTCAGGAGCCACACTCTCCCACTNCCCTTGCCCANAGACTC
ACAGAGGGCACGGAGCTGGCTGTGGTGAGAGGAGGTCCANCAAATTCCTGTCTGCA

Marked-Up Copy

NAAGGGTTCTGAACACCACCGCCTGGCAGCGTGCTGGAGGAGGGATTCTCTTTTC
CTCACAGCAATTCTGACCAGAAACCTGTCAAATCAGGAATGGCTAAAATAAGACCA
GGGTATGAATGACCATCAGCCACAGTAAAACCAAGGCACAGCTCTCCTGAGCCCAC
CCAAGCTGCTGTGGCCCAGACTGGTGACATCACCTCAGGGCAAAAAAAAAA

~~Sequence ID 424~~SEQ ID NO: 86

nt: 420

CGCAGAATGGCTCCCGCAAAGAAGGGTGGCGAGAAGAAAAAGGGCCGTTCTGCCAT
CAACGAAGTGGTAACCCGAGAATACCCATCAACATTCACAAGCGCATCCATGGAG
TGGGCTTCAAGAAGCGTGACCTCGGGCACTCAAAGAGATTCGGAAATTTGCCATG
AAGGAGATGGGAACTCCAGATGTGCGCATTGACACCAGGCTCAACAAAGCTGTCTG
GGCCAAAGGAATAAGGAATGTGCCATACCGAATCCGTGTGCGGCTGTCCAGAAAAC
GTAATGAGGATGAAGATTCACCAAATAAGCTATATACTTTGGTTACCTATGTACCT
GTTACCACTTTCAAAAATCTACAGACAGTCAATGTGGATGAGAACTAATCGCTGAT
CGTCAGATCAAATAAAGTTATAAAATTG

~~Sequence ID 425~~SEQ ID NO: 87

GGAAACTGATGCCAGTCAGAACTCAGATCAAATGAAGGGGTGAAGAGAACCAGAA
TTGATCTCTCTGTAGGAGAATATAAATGACTTTTTTAAAGTACATATTTTCTGTGA
AAGACAGTTTTTTGTTTAATGCAAAAATGTTAACAATGTTTATATCATGTAGAAGT
AAAAGATCGTGAAACAGCACAGAGAACAGTAGTAAGACAGATTGAATTGCACTGTT
GTAAGATGATGAACTTACAATATTAAGTGAAGGTAGACTGTGATAGATTAAAGGATA
TATATTGTAATCCCTAGAGCAATTGTCAAAGTGGTACAGGTAAAAAGCCAATAGAG
GTGATAAAATGGAATACTAAAAAATATCAGATGAATAATAAAGAAGACAGGAAATG
AGGAACAGTGGAACAGAATGAATAAAAAACAAGACCATTA ACTTAAATCATTAAATA
TTACTTTAAATGGGTAAACATTATGGTTATAAGGCAGAGATTTTCAGACTAGATA
AAAGAGCAAGCTCCACTATATACTGTCTACAAGAGATATACTTTAAAGTGTATATT
ATATTTAAATATAAAGATTTGGAATAAATAAACCTAAGAATAAGCTTACTAGGGAA
GTGAAAGATCTGTACAACAAGAATTACAAAACACTGCTGAACGAAATCATAGGTGA
CCA

~~Sequence ID 426~~SEQ ID NO: 88

GTCCCGGAATCGCGGCCGCGTCGACGTTTCCTCAAATTTATCTTCCTGTTAATGT
CAGGCATGTATCTCCTTAGCTTGCCACAAATAACTATATATACCACAGACCTTCCT
TTGTAGGGCTAACAGTGTTGCATTGTAAGTGGAGGCCTCATAGATACCTGGCCTTT
TCCTACCTTATTCCAAAGATGGTTGCATCTTATAAATAATGTCATTCTTCAGCAAA
TGGTATGGAAATGAGATTGTAATGTCATTATTTCTCTTTAAATAATCAGGACAAC

Marked-Up Copy

TCATGATACAAAGAGCTCTTCTCTATAAAAGGTGGGACTTTTTTTTTTTAGTAATAG
CAAAAATAAAATTGTACCTCCTTAATCTTCTACAGAAAGATGGATTTTCATTTTCAA
CATTAAGAGGTAGTTTTAAGAAGCAGTAGAAGTCAGCCTGGGCAGCATGGTGAAAC
CCCGTCTCTACAAAAAAGTTAGCTGGGCTTAGTAGTTGCAATCCCAGCTACTCTGG
AGGCTGAGGTTGGAGATCATCTGANCTGGGGAGGTCNAGGCTGCAATGATACANT
GAGCCCTGATTGTGCCACTCCACCTGGTTGCAGA

~~Sequence ID 427~~SEQ ID NO: 89

TTCCAATCTTCGTGTTCACTTTAAGAACTCGTGAAACTGCTCAGGCCATCAAGG
GTATGCATATACGAAAAGCCACGAAGTATCTGAAAGATGTCACTTTACAGAAACAG
TGTGTACCATTCCGACGTTACAATGGTGGAGTTGGCAGGTGTGCGCAGGCCAAGCA
ATGGGGCTGGACACAAGGTCGGTGGCCCCAAAAGAGTGCTGAATTTTTTGCTGCACA
TGCTTAAAAACGCAGAGAGTAATGCTGAACCTAAGGGTTTAGATGTAGATTCTCTG
GTCATTGAGCATATCCAAGTGAACAAAGCACCTAAGATGCGCCGCCGGACCTACAG
AGCTCATGGTCGGATTAACCCATACATGAGCTCTCCCTGCCACATTGAGATGATCC
TTACGGAAAAGGAACAGATTGTTCTTAAACCAGAAGAGGAGGTTGCCCAGAAGAAA
AAGATATCCCAGAAGAACTGAAGAAACAAAACTTATGGCACGGGAGTAAATTCA
GCATTAAAATAAATGTAATTAAAAG

~~Sequence ID 428~~SEQ ID NO: 90

TGCAGGATCCGTCGACTCTAGATAACATGGCTAGAAAAGAGAATGAAAAAGTTGGA
ATTTTTAATTGCCATGGTATGGGGGGTAATCAGGTTTTCTCTTATACTGCCAACAA
AGAAATTAGAACAGATGACCTTTGCTTGGATGTTTCCAACTTAATGGCCCAGTTA
CAATGCTCAAATGCCACCACCTAAAAGGCAACCAACTCTGGGAGTATGACCCAGTG
AAATTAACCCTGCAGCATGTGAACAGTAATCAGTGCCTGGATAAAGCCACAGAAGA
GGATAGCCAGGTGCCCAGCATTAGAGACTGCAATGGAAGTCGGTCCCAGCAGTGGC
TTCTTCGAAACGTCACCCTGCCAGAAATATTCTGAGACCAAATTT

~~Sequence ID 429~~SEQ ID NO: 91

nt: 535

CACAGTACTCCATTTTGGGGTCCAACTGTAATGCTCAAAATAATAAATGCTTACA
CGAAAATTATTTATTGAGAATATTCATATAAAAATTACCTAAAGCAAAGTAAAAA
AGTAAATCAAGGTGGTATATTTGAAGTGAATGGTGATTGGAAATTTTTAGCTGTA
ACAAAAAGAAAGAAACAACTTTTTTTAAAGCCTCATTCTCTTTTCTTTCAAAATG
TACCTTATTCCCACACACTCTTGGGCTGACCTTTATTTTATCAATAAGCTCAATAT
TACTTTGTTTAAAATAAGATGCTTCAGCAAAAGTCATTCTCTTTTAACCATATAA
TTTAAAACTCCTCTTCACGATTGATAGCAAAATCAGAAACGTTAGGGCACCAGTG

Marked-Up Copy

AGTTGAAAAAACTGGTCTTAAGTTGGAAAACTATTATTAATAATATTATCCTATC
CATCCATATCTATTGAAATTGTCAGGTCATAATTTTCATTTTAATTAATTATAGGA
AAGAAGAAAAGATAATACCCATTTGTTCTAT

~~Sequence ID 430~~SEQ ID NO: 92

CAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAAGTTTTT
TTCTCTTTGAAAGATAGAGATTAATACAACACTACTTAAAAAATATAGTCAATAGGTT
ACTAAGATATTGCTTAGCGTTAAGTTTTTAAACGTAATTTTAATAGCTTAAGATTTT
AAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAAAGGTT
TCTAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAATGTATTT
AAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAGAAGGGC
AATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTTAAAAG
TTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAGAGATTAAACCGAAG
GTGATTAAAAGACCTTGAAATCCATGACGCAGGGAGAATTGCGTCATTTAAAGCCT
AGTTAACGCATTTACTAAACGCAGACGAAAATGGAAAGATTAAATTGGGAGTGGTAG
GATGAAACAATTTGGAGAAGATAGAAGTTTGAAGTGGAAAACCTGGAAGACAGAAGT
ACC

~~Sequence ID 431~~SEQ ID NO: 93

CGCTGGGTGCCTGCAGCGCCTCCCTTGTCTCATATGGTGTGTCCAGCACTCTATTG
TTGTAAACTGTTGNNTTGNCTGACCTAAATTNTCTTTACTAAACANATTTAATAGT
TNAAAAAAAAAAAAANANCA

~~Sequence ID 432~~SEQ ID NO: 94

TTTTAAAGTCATCTCTATAGGAAGGTGCTGGGCAGGGATCCCAGAGAAAGAAAGGG
TCCAAGACTCCATTAACTGCCCTGGATGAAGGGCACTGCTACAGCAGCTAGTACCA
GAGACTCTCCTATCTCACGGTTGAGGCAGACCCAGGATAGAATAGAGAATAAAAGG
AATGCTTATAGGAAACAATTTTGTATGGAATGCTAGATGGCCAAGCCTCAGCCTTT
GGTCCAGTGCAACCCTTGCCCTCGCTTGTCAACAGTGAAAAATTAGTTTGGTTAGAA
GAACCATCTGGAAACACACCAGCTTCTGCTACCTTCATGCTCATTGTTAAAAAAG
ATTAACCAGTGTGAACATTCTGATCTGTTAATTCCAGGGACTGTTTTCTTTCCAAT
GGACTGTTTGTGTTGGTAGAATAACCCCCAAAAGCTCAAAGCTAAAATGCATCATCAG
TCCTAGTCGGCAGTTCCTTAAGAATGGACTGGCGGCGTGGTTGAGCTGATATGGAA
AAGCTGCACCTTCCTGCAGAAGATCAACTGACCTGCTATCCCACCCCAAATTTCAA
CCTGAGGTATATTTCAATGAAGGCAGGTAGCTGTGCTTCTCAGAGCA

~~Sequence ID 433~~ SEQ ID NO: 95

TCCCGGAATCGCGGCCGCGTCGACCCGCCGCCGAGGATTCAGCAGCCTCCCCCTTG
AGCCCCCTCGCTTCCCGACGTTCCGTTCCCCCCTGCCCCGCTTCTCCCGCCACCGC
CGCCGCCGCTTCCGCAGGCCGTTTCCACCGAGGAAAAGGAATCGTATCGTATGTC
CGCTATCCAGAACCTCCACTCTTTTCGACCCCTTTGCTGATGCAAGTAAGGGTGATG
ACCTGCTTCCTGCTGGCACTGAGGATTATATCCATATAAGAATTCAACAGAGAAAC
GGCAGGAAGACCCTTACTACTGTCCAAGGGATCGCTGATGATTACGATAAAAAGAA
ACTAGTGAAGGCGTTTAAAGAAAAAGTTTGCCTGCAATGGTACTGTAATTGAGCATC
CGGAATATGGAGAAGTAATTCAGCTACAGGGTGACCAACGCAAGAACATATGCCAG
TTCCTCGTAGAGATTGGACTGGCTAAGGACGATCAGCTGAAGGTTTCATGGGTTTTTA
AGTGCTTGTGGCTCACTGAAGCTTAAAGTGAGGATTTTCCTTGCAATGAGTAGAATTT
CCCTTCCTCCCTTGTCACAGGTTTAAAAACCTCACAGCTTGTATAATGTAACCATTT
TGGGGTCCGCTTTTAACTTGGACTAGTGTAACNTTCATGCAATAAACTGAAAAG
ACCATGCTGCTANTC

~~Sequence ID 434~~ SEQ ID NO: 96

TTCGGACGCAAGAAGACAGCGACAGCTGTGGCGCACTGCAAACGCGGCAATGGTCT
CATCAAGGTGAACGGGCGGCCCTGGAGATGATTGAGCCGCGCACGCTACAGTACA
AGCTGCTGGAGCCAGTTCTGCTTCTCGGCAAGGAGCGATTTGCTGGTGTAGACATC
CGTGTCCTGTAAAGGGTGGTGGTCACGTGGCCANATTTATGCTATCCGTCAGTC
CATCTCCAAAGCCCTGGTGGCCTATTACCANAAATATGTGGATGAGGCTTCCAAGA
AGGAGATCAAAGACATCCTCATCCAGTATGACCGGACCCTGCTGGTAGCTGACCCT
CGTCGCTGCGAGTCCAAAAAGTTTGGAGGCCCTGGTGCCGCGCTCGCTACCAGAA
ATCCTACCGATAAGCCCATCGTGACTCAAACTCACTTGTATAATAAACAGTTTTTT
GAGGGATTTTAAAA

~~Sequence ID 435~~ SEQ ID NO: 97

CTGCAATGTGCAATAGTTGCACCACTGCACTCCAGCCTGGGTGACAGAGTGAGAAC
CTATCTCTTAAAAAAAAAAAAAAAAAAGGAAGAAGAGACATGAGAGGGCCCAAGT
CACTTGCTCACTCACTTTCCGTGTACATGTACCAAGAAAAGGCCATGTGGGAAAGA
GCAAGAAGGCAGCCGCTTCAAGACAGGAAGAGAGCCCTCACCAGAACTGAGCCA
GAACCTTGGAATTCAGCCTCCANAACGTGTGAGAAAAGAATTTTCTGTTGTTTCAG
TCCCCACACTATGGCATTTTGTTACGGCAGCCTGAGCTAATACTCCTACTTTGTC
CTGCATTTACTTGGTCTTCCAGTTAGTTTTTTAGACTTTGGGAATCAGAGCAGTCA
GTTGTCAGATTTTAGCTTACAGTTGTCCTACCTGTGCAACTGAAATTTCTTCCATT
TTAAACCAGAGCAGAGTTTTAGAGTCAAAGAAACCAGATCTTTTAGTGCAGAAGC

TTTCCACTGTATTANAAGTGAGGAAGTTGGT

~~Sequence ID 436~~ SEQ ID NO: 98

AAAAAACTCCAGAGAAGTTTATAGAAAGAGATGACATGTAAACCCTGCTGAAAAA
TAGTTTCATTTGTTAGAATATAATTGTCTTCCACTAAAAAAGAAAAAAAAAAGCA
TTTAAGGCTCTAAGATCTCTTGAAGTACCACTTTTCCTGAATCCCAGAGTTTTTAT
GTGCATTATTTTTATGCGTTTGTAGTTTGATATGTTGTATTTATAAGTAGTTTTAG
CTTTCCATTATGAATTCTTCTTTGACCCATGAGTTATTTAGGTAAGTGTTTAAAAA
TTTACAATAGTTTATATATGCAAATATTATGTTGTTAGAGTTGGTTTTTCATGTCAT
TTTTACATATACAGGGGCAGTTTCCCCAACTAAATTGTATATTCCTTAAAGCAGCA
CTCTTAAATTTTATTTCTGTGTCAATTTCTTGNCTGTGTTTCCTGGCATGGAATAC
ATGGCATAAAATTTGTTATGTAATTAAATGAAATATTATTATACTTTCTATTTTTT
AGAAAAAA

~~Sequence ID 438~~ SEQ ID NO: 99

nt: 577

GTCGACAGGGATGACATAACTATTAGTGGCAGGTTAGTTGTTGGTCACTTTCAACT
CTGGGTTCAAGCGATTCTCCTACCTCAGCCTCCCGAGTAGCTGGGATTACAGGCAT
GCACCGCCACACCTAATTTTCTATTCTTAGTAGAGACGGGGTTTCTCCCTGTTGGT
CAGGCTGGTCTCGAACTCCCGACCTCAGGTGATCTGCCTGCCTCAGTCTCCCAAAG
TCCTGGAACCACAGACATGAGCCACCACGCCTGGCCCCTTTTAAATATTTCTGCT
CATTGATGATGCACCCAGTCACCCAAGTGCTCTGATGGAGATGTATAAGGAGATGA
ATGCTGTTTTTCATGGCTGCTAATACAACATTCATTCTGCAACCCCCAAATCAAGAA
GTAATTTTGACTTTCAAGTCTTATTATTTAAGAAATATATTTTGCAAGACTATAGC
TGCCATAGACCGTGATTCTCTGATGGATCAGACAACTAAAATGAAAACCTCCTG
CAACGTATTCATCATTCTAGATCCCTGAGGAATCGCCACACTGACTTNCACAATGG
GTGAACTGGGTTACAGT

~~Sequence ID 441~~ SEQ ID NO: 100

nt: 552

AAACAAAATTATTCTCTGAGAGGGAAAGGACATTTGAGGGAAACATCAAATTTCCC
CATAAATAAATGAATGGAGTTTGCAGGAAGGTGAGGGTGAGCAGAGATGTGTGTGG
ACATCTCTGACCATCCATCGCTGTATTCAAATGGATTGTTTTATTCCATTCTGGTC
TCAGGCATGACCACGTCCAGTGAAGACATTTGAGGCAGCACATCTCAGGACCCAGG
CAATAGACTGGCCCCAACTCAGGCTGGACTAAGGTGTGATTAATTCTTTGTTTTTT
GTGTGGAACAGCTCACCTTGTCAGACAGCCTCAGGGCATCTCTGAGACACAGGGGC
AGAAAATGACATTCATCTTTTGAGTCCTCATCCATGGAGTGCTGTGTTTGGGGGGC

~~Sequence ID - 442~~SEQ ID NO: 101

~~Sequence ID 446~~SEQ ID NO: 102

~~Sequence ID 447~~SEQ ID NO: 103

AGGCCGAGGCAGGCAGATCNCNTGAGGTCAAGAGTTTGAGACCAGCNTAGCTAACATGGTGAAACCCCATCTCTACAAAAATATA-AAAATTAGCCTGG-GTGGTGATGGGGACCTGTAACCCAGCTACTCGGGAGGCTGAGGTAGGAGAATCACTTGAACCCGGGAGATGGAGGTTGCAGTGAGCCAAGATCGTGCCACTGCACTCCAGCCTGTGTGACAGACAAGACTCTGTCTCAAAAAAAAAATAATAATAATAATAATAAAAAAGGAATAACATAGCTAGGAATAAATTTAATCAAAGAGGTGAAAGACTTATACACTTAAACTACAAAAAAAATCACTGAAGGAATTATAGACCCAAATAAAAATAAATAAAAAGACATTCTGTGTTTTAGGGAAAGAAGACTTAATATTGTTAAGATGTCAATACTACCCAAAGTGATCTACAGATTCAACATAATCCCTATCAAAATTCCAACAGCCTACTTTGTAGAAA

Marked-Up Copy

TGGAAAAGCCAATTTTCAAATTCAGATGGAATTGCGAGGGGTNTGAATAACAAAA
CACNATCTTGGGGAAAAAACAACAAAGTCAAAGAACTCACACTTCTNTAT
TTATAAATTTACTACAAAGTTATAGTAATCNAA

~~Sequence ID - 448~~ SEQ ID NO: 104 nt: 329

TACGCACACGAGAACATGCCTCTCGCAAAGGATCTCCTTCATCCCTCTCCAGAAGA
GGAGAAGAGGAAACACAAGAAGAAACGCCTGGTGCAGAGCCCCAATTCCTACTTCA
TGGATGTGAAATGCCCAGGATGCTATAAAATCACACGGTCTTTAGCCATGCACAA
ACGGTAGTTTTGTGTGTTGGCTGCTCCACTGTCCTCTGCCAGCCTACAGGAGGAAA
AGCAAGGCTTACAGAAGGATGTTCCCTCAGGAGGAAGCAGCACTAAAAGCACTCTG
AGTCAAGATGAGTGGGAAACCATCTCAATAAACACATTTTGGGTAAAA

~~Sequence ID 450~~ SEQ ID NO: 105

GAGCAGTGGCATGATCACACCTTACTGCGGCCTCCAACCCCTGAGCTTAAGTGATT
CTCCCGCATTATCCTCCTGAGTAGCTGAGACTACAGGTGCATGCCACCATACACTA
CTAAATTTGGGTGCGGTGGTGGTGGTGATTTTTTAATATTTTGTAGAGACAGGGT
CTCACTGTGATGCCCAGGCTGGTCTTGAACCTCTGGGCTCAAGCAGTCACCCACCT
CAGCCTCCCAAAGCACTGGGATTACAGGTGTGAGCCACCACACTGGCCAGCTTTGT
TTTGTTTTGATGACTAAGCTGCTCTTGCTAAAAGGGCTTCTCTCTGAACTTCCCTA
CCTTTCTTCTGTTTCCCTGGGCTAGGGCTCCATGTTGGCAGTCCTACTCCCAATTA
ACCTGGGGCTGTCTGGTTAACCTTTATAAGATCTGCAGTCATTGGGAGACCCGGGG
ACCAGGAATATTGTTGTTGAGGGAGCTACCCTGGAAAGTGGATGGGTGGCCAAAGG

~~Sequence ID 452~~ SEQ ID NO: 106

TTTGGCTTTGCCTCTAGGCATTAGATGTTATCTTTGGAGGCATCCTTCTATGAGCA
TTCATTTTTGGACCAAGCCTGGATTTACAATTCTATTACTGGCCCAGACTTCATTT
CTATCCAATTTCAATCCACTGTGCTATAGTTTACAACATATAATTTGACTTATAAA
TAATTCCTGACTATGGGTTTTAAAGACTGAAAATGGATCAATAGAACTTTGAAAAT
GTTAACATCTTGATTGCTTTTCTCAGTGTAGAAATGGACAATGTTTAGCTTAAAAA
CTGCATGTTTTTAATGAGATACGGGGTTGAAAGACTTATTCCTGGAATTTATTGTT
CTGGAGAAAGCCTGTTGCTATCTGCCATACCTTGGTTTACTTTGTGCAAAATGAGC
TTCTTTTTAAGTAATGAGCTCTTTCCATGTTTCAGCTTAAATTGCTGTCTTAGACAC
TTCATCAGGGTTCCTGCTCTGCCTCATTCCTTTTGGCTCACTTGCAGCCTTTG
ACATAATCCTGGGAGGCAATTGGCATCATACATATTTTGCTTTGTAATCTCCTGCT
TTGATTCTGACTGGGACCCAGC

Marked-Up Copy

~~Sequence ID 453~~ SEQ ID NO: 107

nt: 747

GGATCTAAGACCAGCCTGGCAGCCACCAGATGGTGATTCTAGTCCTGGCTCAGTCA
GTAATAGGTCACTGACCCCAGAGAAATCAATTCAGCCTCCCCAGGTCCTTGGATTT
CTTTCTGTGAAAATGAAAGCATAGGTAGGAATTTCCCATGGAACAGCTAGCAGAGG
AGAAATATTAAAAGTCAGGAGACTCATGCTATAGTTTTTCATACTTCATTACAACAA
TGTTGTTTAGGACAAGTGAGTTAACCTGTTAGCTTCCTCTATATAAAATGGAAAGT
CATTAAAAACCTACATAGCAGGGTTCTTGTGAAGATCAAGTGATAATGTAGGAAGC
ATGTACAAATGTCACATTCTGCCGTCACGTAATGGTCCTCACAGCTTGAGGTAGCA
TTTAGCATGTGTCATGATTTAGTACAAGGGTTGGCAAACCTGTTGCTCTTGGATTAA
GTCTGGCTCATTGCCTGTTTTTCAAAGAAAAAAATTGTATATGTGTGTATATATGT
TATATATAGGTACACACACATATGTGCTATATATAGCATATATACACACATAATAT
ATAAACATGTACATATATAGCATTATATATATATACCGTGTATAATATCTCCAGTCCT
CATGACCAGCCATGCTTGTTTACATTTGCATACTCTATGATTGCTTTTCATGC
AACAATGGCAGAGTTGAGTGATTGTTTTGCACAGANACTGTATGGCCCACTAAACC
TAAAATATTAATCTCTGCC

~~Sequence ID 454~~ SEQ ID NO: 108

CTCCTGCCGGGCTCGTGGCGGCTTCTGTCCGCTCCGCGGAGGGAAGCGCCTTCCCC
ACAGGACATCAATGCAAGCTTGAATAAGAAAAACAAATTCCTCCTCCTAAGCCATG
GCATATCAGTTATACAGAAATACTACTTTGGGAAACAGTCTTCAGGAGAGCCTAGA
TGAGCTCATAAGTCTCAACAGATCACCCCCCAACTTGCCCTTCAAGTTCTACTTC
AGTTTGATAAGGCTATAAATGCAGCACTGGCTCAGAGGGTCAGGAACAGAGTCAAT
TTCAGGGGCTCTCTAAATACGTACAGATTCTGCGATAATGTGTGGACTTTTGTACT
GAATGATGTTGAATTCAGAGAGGTGACAGAACTTATTAAAGTGGATAAAGTGAAAA
TTGTAGCCTGTGATGGTAAAAATACTGGCTCCAATACTACAGAATGAATAGAAAA
ATATGACTTTTTTACACCATCTTCTGTTATTCATTGCTTTTGAAGAGAAGCATAGA
AGAGACTTTTTTATTTATT

~~Sequence ID 458~~ SEQ ID NO: 109

nt: 682

TGCCACTGAAGATCCTGGTGTCGCCATGGGCGCGCCCGCCCGCTTGTACCGGT
ATTGTAAGAACAAGCCGTACCCAAAGTCTCGCTTCTGCCGAGGTGTCCCTGATGCC
AAGATTCGCATTTTTTGACCTGGGGCGGAAAAAGGCAAAAGTGGATGAGTTTCCGCT
TTGTGGCCACATGGTGTGAGATGAATATGAGCAGCTGTCCTCTGAAGCCCTGGAGG
CTGCCCCGAATTTGTGCCAATAAGTACATGGTAAAAAGTTGTGGCAAAGATGGCTTC
CATATCCGGGTGCGGCTCCACCCCTTCCACGTCATCCGCATCAACAAGATGTTGTC
CTGTGCTGGGGCTGACAGGCTCCAAACAGGCATGCGAGGTGCCTTTGGAAAGCCCC

Marked-Up Copy

AGGGCACTGTGGCCAGGGTTCACATTGGCCAAGTTATCATGTCCATCCGCACCAAG
CTGCAGAACAAAGGAGCATGTGATTGAGGCCCTGCGCAGGGCCAAGTTCAAGTTTCT
GGCCGCAGAAGATCCACATCTCAAAGAAGTGGGGCTTCACCAAGTTCAATGCTGAT
GAATTTGAAGACATGGTGGCTGAAAAGCGGCTCATCCCANATGGCTGTGGGGTCAA
GTACATCCCCAATCGTGGCCCTCTGGACAAGTGGCGGCCCTGCACTCATGAAGGCT
TTCAATGTGC

~~Sequence ID 459~~ SEQ ID NO: 110

TCCCGGAATCGCGGCCGCTCGACCTTGTCTTGAGCGTCAACCTTCTTTCCTGA
AGTGGCTGGGGTTCCTGTTTCCTTCTTTGATTGACAACTTGTGTTAACCTTCGCAC
ATCTCTGGGCCAATTTTTGCTTGTAAGTCTTTCGGAGACCCCTGGAATTTAAATC
ATTAGCACCGCGCCCTTCCCCGAAGAGTCTTCGAAGGGTTGCCGCTTTTCGGTGGC
GCAGTTCTCGCGAGAAGGTGACTTTCTTCTCGGTATTTCTCTGGTTTCCAGAATCC
TTAGCGCGAGGCGGAAAAAATATTTCTCCAGCTTGTGTTGATGCCGCGATTTTGA
CTGAGACTTCTTCCCACGATTTCTGTTTTTGCTTCTCCAAGGAAAAATGGCAGCTCC
CGAGCAGCCGCTTGCGATATCAAGGGGATGCACGAGCTCCTCCTCGCTTTCCCCGC
CTCGGGGCGACCGAACCTTCTGGTCAGGCACCTGCCGGCTGAGCTTACTGCTGAG
GAGAAAGAGGACTTGCTGAAGTACTTCGGGGCTCAGTCTGTGCGGGTCTGTGTCAGA
TAAGGGGCGACTGAAACATACAGCTTTTGCCACATTCCCTAATGAAAAGCAGCTN
TAAAGGCATTGACAACTNCATCAACTGAACTTTTAGTCATACTTTAATCG

~~Sequence ID 460~~ SEQ ID NO: 111

nt: 536

CAGAGATCAAAATAGGCCTTACACAGTGCAGCGAATTTAAAAGATTACCCCAT
CAGGTGTATGGATTTTGCAGTATTAAAGATGCTGCCTGGAATAGGTCATTATCTTC
TCCAAGTACTCTGTAAAGTCAATGAGTCACATAGAGTATAAGGTTTATTATCTGCT
TTTCTTTTCATTAAATAAATCTTTATTGAATTTCTACTACATTA AAAA ACCAAACCA
AAACAAAACAAACAAAAAAACACTTCCCTGAGCCATAAAGGAGAAGGTAGTTTTG
ACTGGAACCTTGAAGGATGGGTAACTTTTCAGCAGATAAAGATTGAGAGAAGACCT
TCCAGGTAGAGAAAGCAGTGTGGGCACAGGCAAAGATGGAAGAACACACGTGGCTG
TGGGAAACACAGCTAGAAGCCAGTGCGGATAGAGAGTAGGCTATGATGTGCAAAGG
TTANACACTGGGAGAGACAGGTCCATGAGAGTAGCTTGGACTAACACAGGGAGGGT
TTGGAATCCCAACTGGGGAACCTANAAATCAA

~~Sequence ID 461~~ SEQ ID NO: 112

TAGGAGGCTTATTCAGTATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCT
ACGCCAAAATCCATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCA

Marked-Up Copy

CAACACTTTCTCGGCCTATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGC
ATACACCACATGAAACATCCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAG
TAATATTAATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAAGCGAAAAGTCCTA
ATAGTAGAAGAACCCTCCATAAACCTGGAGTGA CTATATGGATGCCCCCACCCTA
CCACACATTCGAAGAACCCGTATACATAAAAT

~~Sequence ID 462~~SEQ ID NO: 113

TCTTTATCAAGTTGAGAAAGTTCCTCCCCTCTATTCCTAGTTTGCTAAGAGTCCTT
CTATCCTATTTCTTAATGGTTTAGTAGATGACTCTGTGGTACTTTGAAGGTTGTTT
GCAGAATTTCCATGCCATAGGCAATTTACCTTTCCTTGACATTTGAAGGATTGATG
TTGGTGCCAAGTATAGAATCTTCACAGAGTCCTCCTGTAGCTTCTAAAGGTTTAGC
TTGAAAATGTTAATTGCTTAACGCTAGTAAGTGAGTGAAAAAGCTGGGGATAAATT
TTGTATCTTGCTTATATTTTCAGTTCCACCTCTGTCCNGACNAAACCCCATATAT
AA

~~Sequence ID 463~~SEQ ID NO: 114

TAGTTTACATATCCCAACCTTTAAAAATATTCCTCTTATTAGCTTTATATTTACTT
TATAGAAGTTGAGTTTAAATTAATAATTCCTGGCATCCTGAAGTATGTCACATAGCA
TGTGCTCCTTATAAATATGTTGATATCTCAGAAGACAGCATCCCGGTTTTTCATTTT
ATAAAGTACCATACTTAAGAATGCTGTAATACTTATCTTTTATAACATGTTTCCTT
CGCTTTGCTTGNCTTTTATGNCATCAGTTTAACTGTTTACTTTCATTTAACAGNTT
ACATCATNCAACAGTTTACTTCATTAAACAGTAGGTGGAAAAATAGATGCCAGTCT
ATGAAAATCTTCCCATCTATATCAAATACTTTCAAGGATATACTTT

~~Sequence ID 464~~SEQ ID NO: 115

nt:

615

CGACTTTCAACCATCAAGTGAGGAATACCTTCACATAACTGAGCCTCCCTCTTTAT
CTCCTGACACAAAATTAGAACCTTCAGAAGATGATGGTAAACCTGAGTTATTAGAA
GAAATGGAAGCTTCTCCACAGAACTTATTGCTGTGGAAGGAACTGAGATTCTCCA
AGATTTCCAAAACAAAACCTATGGTCAAGTTTCTGGAGAAGCAATCAAGATGTTTC
CCACCATTAAACACCTGAGGCTGGAAGTGTATTACAACCTGCCGATGAAATTGAA
TTAGAAGGTGCTACACAGTGGCCACACTCTACTTCTGCTTCTGCCACCTATGGGGT
CGAGGCAGGTGTGGTGCCTTGGCTAAGTCCACAGACTTCTGAGAGGCCACGCTTT
CTTCTTCTCCAGAAATAAACCTGAAACTCAAGCAGCTTTAATCAGAGGGCAGGAT
TCCACGATAGCAGCATCAGAACAGCAAGTGGCAGCGAGAATTCTTGATTCCAATGA
TCAGGCAACAGTAAACCCTGTGGAATTTAATACTGAGGGTGCAACACCCCATTTTC

Marked-Up Copy

CCTTCTGGAGACTTCTAATGAAACANATTTCTGATTGGCATTAAATGAANAGTCA

Sequence ID 469

GATTTTTTAAAAATACATATAGCAAAAATATTACAGGGTCAGGGGAGACAATTAGAA
TGATATAATTCAAAGTGGATTAAAAAAAATGTCACCCAGAATACAATACCCAG
CAAAGTTGTCCTTCATAAATGAAAGAAAAATNAAATCTTTNCCNAACNA

~~Sequence ID 471~~SEQ ID NO: 117

TCCCGGGAATCTGCAGGATCCGTCGACT

~~Sequence ID 472~~SEQ ID NO: 118

GACAGTGCCCAGGGCTCTGATATGTCTNTCACANCTTGNAAGTGTGAGACAGCTG
CCTTGTGTGGGACTGAAAGGCAAGATTTGTTCTGCCCTTCCCTTTGTGACTTGAA
GAACCCTGACTTTGTTTCTGCAAAGGCACCTGCATGTGTCTGTGTTCTTGTAGGCA
TAATGTGAGGAGGTGGGGANACCACCCACCCCATGTCCACCATGACCCTCTTNC
CACNCTNACCTGTGCTCCCTCCCCAATCATNTTT

~~Sequence ID 473~~SEQ ID NO: 119

nt:

694

TGGGCTTTGGGCTGGCTGCAGTCTGTCTGAGGGCGGCCGAAGTGGCTGGCTCATTT
AAGATGAGGCTTCTGCTGCTTCTCCTAGNGGCGGCGTCTGCGATGGTCCGGAGCGA
GGCCTCGGCCAATCTGGGCGGCGTGCCCAGCAAGAGATTAAAGATGCAGTACGCCA
CGGGGCCGCTGCTCAAGTTCCAGATTTGTGTTTCCTGAGGTTATAGGCGGGTGTTT
GAGGAGTACATGCGGGTTATTAGCCAGCGGTACCCAGACATCCGCATTGAAGGAGA
GAATTACCTCCCTCAACCAATATATAGACACATAGCATCTTTCCTGTCAGTCTTCA
AACTAGTATTAATAGGCTTAATAATTGTTGGCAAGGATCCTTTTGCTTTCTTTGGC
ATGCAAGCTCCTAGCATCTGGCAGTGGGGCCAAGAAAATAAGGTTTATGCATGTAT
GATGGTTTTCTTCTTGAGCAACATGATTGAGAACCAGTGTATGTCAACAGGTGCAT
TTGAGATAACTTTAAATGATGTACCTGTGTGGTCTAAGCTGGAATCTGGTCACCTT
CCATCCATGCAACAACCTTGTTCAAATTCCTGACAATGAAATGAACTCAATGTGCA
TATGGGATTCAATCCCCACCATCGATCATAGCACCCCTATCAGCACTGNAACTC
TTTTGCATTAAGGGATCATTGC

~~Sequence ID 474~~SEQ ID NO: 120

GGCAGCGCGGGGAGCCCGTCGGCGCCGGCGGGCGGGCCGGTTTCGAAGTTGATGCA
ATCGGTTTAAACATGGCTGAACGCGTGTGTACACGGGACTGACGCAACCCACGTGT
AACTGTCAGCCGGGCCCTGAGTAATCGCTTAAAGATGTTCCCTACGGGCTTGTTGCT

Marked-Up Copy

GTTGATGTTTTGTTTTGTTTTGTTTTTGGTCTTTTTTTGTATTATAAAAAATAAT
CTATTTCTATGAGAAAAGAGGCGTCTGTATATTTTGGGAATCTTTTCCGTTTCAAG
CATTAAGAACACTTTTAAATAAACTTTTTTTTGATAATGGTTAAAAAAAAAAAAAAAA
A

~~Sequence ID 475~~ SEQ ID NO: 121

CATAATAAAAAACAATCAACAAACAGGGAATGGAAAGAACTTCCTCAGCATGGTG
AAGGCCACATATGAAAATCCACAGCTAACATCATACTCAATGATGAAAGACTGAA
AGCTTTTCTCCTGAGATCAGGAACAAGACAAAGATGTCACCTTTTGTCACTTCTAT
TCAACTCATTATTGGAAGTTTTTGCCAGAGCAATTAGGTAAG

~~Sequence ID 476~~ SEQ ID NO: 122

nt:

476

CAGAATCTTTTCATAGGCTGAATGTTGCTCCACAATGTGTCCTTTGACTATCTCTG
GCTAATTATTATTTTAATCTCTTCTCAGCTTTTCCAAGAACATAACGTTAACCAAA
GATCTTAGGCCATTCACAACTCTTTTGTAATAAATTAATGTGGATGTGAAACGAGGC
AACAAATCCTGAAGTAGAAAGTTATTCCTGGCCAGGCACGGTGGCTCACGCCTGTA
ATCCTGGCACTTTGGGAGGCCGAGGTGGGTGGATCATGAGGACAGGAGATCGAGAC
CATCCTGGCCAACATGATGAAACCCCATCTCTACTAAAATACAAAAAATTAGCTGG
GCATGGTGACGCGTGCCTGTAGTCCCAGTTACTCGGGAGGCTGAGGCAGGGGAATT
GCTTGAACCTCGGAGGTGGGAGGTTGCAGTGTGCCGAGATCACGCTACTGCACTCC
AGCCTGGCAACAGAGCAAGACTCCATCT

~~Sequence ID 477~~ SEQ ID NO: 123

AAACAGAAAGTTTCTTCTAAAGGCATGATTCAGTTAAGTCATTCTTAAGTGTTAAA
AAATTGTGAAAAATGTGCCTGTAATCCCAACACTTTGGGAGGCCGAGGCAGGCAGA
TCACGAGGTCAGGAGATCAAGACCATCCTGGCTAACAAGGTGAAACCCCGTCTCTA
CGAAAAATACCAAAAACATTAGCCGGGCGTGGTTGTGGGCGCCTGTAGTCCCAGCT
ACTTGAGAGGCTGAGGCAGGAGAATG

~~Sequence ID 478~~ SEQ ID NO: 124

TTCTTGGGATATTGATGACTACTGTCTGAGAGGTGCTGTGGGGAGATTTTCAGGAT
TGTGTGGTCTTTGAGGGGGGTGTTTTTTTAAGACAACATTGACCACTGTCCACTGT
CCACATGATCATTTGTAATAATTGCAATGCCGCATGCTAGTTGGTTACATAAGACATA
ATTCCAGTGATTGAAGGTGGTTACACTGTATGGTGGTGTGTTCAAGATGGCACTGG

Marked-Up Copy

CATCTTTGAGCAGAGCCTGGCTATGCAGCATCATTTGAGTTTTTTTAAACACCCTAN
AGGTCTGGTTGTTGTTGCTGTTGTCCTTTCCTGTGAAAGTCACAAANANAAGTTACA
GTCCAGGTGAACCTGGAGTTTATAGGTTGGTTTTGTTTCTGNTATATATATATATA
TATATATTTTTTTTTTTTTTTTTTAAACATTTACCTGTAGTGCTGTAGCTGTTGATACTA
TCACCTGCATGCTATTTCTAGTGAGTGCTAAATACAGTATGGTCCAATGACAATAA
CAGCCCATGGTACTGCCAG

~~Sequence ID 479~~ SEQ ID NO: 125

CATCAGTCTGTTATCCATGCTGACTTTCGAAGACTTGCAGCTACTGCATTGATAT
CTTTCCTGCCAATAAGCAAAGTGTTGAACACTTCACAAAATATTTTACTGAGGCAG
GCTTGAAAGAGCTTTCAGAATATGTTTCGGAATCAGCAAACCATCGGAGCTCGTAAG
GAGCTCCAGAAAGAACTTCAAGAACAGATGTCCCGTGGTGATCCATTTAAGGATAT
AATTTTATATGTCAAGGAGGAGATGAAAAAAACAACATCCCAGAGCCAGTTGTCA
TCGGAATAGTCTGGTCAAGTGTAATGAGCACTGTGGAATGGAACAAAAAAGAGGAG
CTTGTAGCAGAGCAAGCCATCAAGCACTTGAAGCAATACAGCCCTCTACTTGCTGC
CTTTACTACTCAAGGTCAGTCTGAGCTGACTCTGTTACTGAAGATTAGGGAGTATT
GCTATGACAACATTCATTTTCATGAAAGCCTTCCANAAAA

~~Sequence ID 481~~ SEQ ID NO: 126

CACACTTTCATGATAAAAACAGAACCTAGGAATGAAAAGAAATTATAGCAACATAA
TAAAGACCATATATGAGAAGCCCACAGCTAACATACTGTATGGTGAAAACTGAAA
GCTCTTCCTCTAAGATCAGGAACAAGGCAAGGATGCCCATTCCTTGCCACTTCTATC
GAACGTAGTACTGGAAGCCCTAGCCAGAACAAC TAGGCAATAGAAAGAAATTAAAG
GCATCCATNTCAGAAAGGAAGANCAAAATGCTGTCTGTTTAAATGACA

~~Sequence ID 482~~ SEQ ID NO: 127

TTTCTATANAAAAAAATTTTTTAAAATAATTGTAAAGTTAGATTTAAAATTGTAAA
ATATAAAATCACAAAGGAATGTACCCAATAAAATGTAAATGCNCCATAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

~~Sequence ID 483~~ SEQ ID NO: 128

CGNTAACGTGCAATCCGCCGCACGCCAGCAAACCTGGACAAACTCCGGGATCTCATC
GAAGCGATTGAGCACCAGTACCAGAGTAATACCGGACTGATGTAACGAGGCGAGTC
GCTCATCCAGCTTGCTGACGTGAGGCAACATCCAGGCCATCGAACGGNTCATCAAG
AATCAACAAGTCAGGCTCCGACATCAGCGCCTGACACAGCAGGGTTTTTCGCGTCT
CGCCAGTGGAAGGTATTTAAAGCGTCNGTCGAGGAGGGCGGTAATACCGAACTGC

Marked-Up Copy

TGCGCCAGTTGCATGCAACGCGGTGCATCCTTTACTTCATCCTGAATGATCTCAGC
CGTAGTGCGTCCGGTGCCATCTTCGCCAGGGCCGAGCATATCGGTGTTATTCCGCT
GCCATTTCGTCGCTGACGAGTTTTTTGCAATTGCTCGAAGGAGAGACGAGTGATGTGG
GAAAACCTGGCTTTGCCGTTACCTTTCAAAAGCGGGAAGTTCCCCCGCCAGCGCGC
GGGCCAGGGCCCGAT

~~Sequence ID 484~~SEQ ID NO: 129

TTTTTTTTTTTTTATTCTATTAATAAATGTTNNTGAAAAAGATACTTAAATTTTAA
AGATAACTNAATTCCTAANGATTTAAATAATCCAAGCAGAGATGAAAGANCAAAT
GCAAATGCNTAAAAAGACCCCANAGCATTGTTAGCAAAAAGCAAATATAGTTAGCC
AAGCATATATATNTCATAAAAGCAATAANAAGGCNTAAAGCAAGTTTGGGGAGAGC
TTATTTAAAACTTGTAATAAATCATTTGAATTTTTTAAAGTTTTCAAAC

~~Sequence ID 485~~SEQ ID NO: 130

nt:

551

TTTGGAACACAAAGTTCCCTTTTTAGAAAGATAGGTATTGAGCCCTTGAGCGTGGG
TAGAAAGATAGAGACAGAGTGATTTGCAAAATAATGGAGGATCATATTTATATATG
AATTTTCACTTATTTGAACTTTCAGATATCANCTTNAAAANCTTTGGTTTAAGTAA
AGTNTNTTAATGAGACTCCTTGGATGAAAGTAACCAAAACCAGTAAAAATAAGGTA
ATAAGGATGTAATAGTTTCTTATGGACACTCAACAGCTAGAATGCAGTTAGTCTCA
GAAAAGAATTAGAACAAATAACTGGAAGGCCATCAGGAGTCCAAAACCATCACTCT
TTTATATTTTATATTTTATTTTCTCTCTTCANATGAGCATTCTCTTTCTATGTCC
ATATGGGTANAAGGCGGCAGCTCCATAGATTATGGCTTCAGATGTTACAGTTCCGCT
NAATGCAGGGACAGACTTGCTATCTTTCAGTCCCCTTACATATCCTGGGGAGAGAG
CAAATGATTGACTGGCTTGAGTCAGGTGCCCGTTCCCTTTCCAATCT

~~Sequence ID 487~~SEQ ID NO: 131 nt:224

GTTTGNTTGTGACCATCTGTACTTGTAATTTCTTTACNTTCATTGGTATGAAAAAT
ATGTTCTTAGAAGCANGAAAAAGAATTCAGNTTTGCTTTGTATACTAAATTAAATG
CTGTAATTTTGATAAAATGAAAAATCTGCTTTATTTGCAACAATTGGTTTCTTCCT
TGACGTCAGCCTCACTCTTGGACTTTGGTATTCAGCCNGNCACCCCTGGGAATTCC

~~Sequence ID 488~~SEQ ID NO: 132

nt:

349

GTGCCTCCCTGTGTGAGTAGCCTAAGGTGCATTGAAAAAGACTGGGATGTGTTTTA
TTTTTTTGTATTAGATAGCATTAACCTTACTGTTGAAGTATTTTTGGTGGAGTATT

Marked-Up Copy

AGTGACAAGCCATTGAGTCTTAAGCCTTACGGCTTCCTATAAAATCACTAATTTTCG
TGTGTGTTTGTGTGTAGGTTACGTTATATATAGGATTCGTGTTCCGCGTGGTGGCC
GAAAACGCCCAGTTCCTAAGGGTGCAACTTACGGCAAGCCTGTCCATCATGGTGT
AACCAGCTAAAGTTTGCTCGAAGCCTTCAGTCCGTTGCAGAGGANCGAGCTGGACN
CCCTGGGGGGGCTC

~~Sequence ID 489~~ SEQ ID NO: 133

TTAACAGCTGCATAGAGTTTTTAAAGTACATTATATTTTGTCTCAGACAAGTAAAATA
TCTGTTTTTTCACGCAAAAAAAGCCATGAAATACGTAATTTTTTAAAGACAAAAAT
CATCTTTTGAGTTTGCTCTTTGGTTTTTCTTCATTCCTTTTGAGGATTGGGAAAAC
AGAAAGATTCTTTGATTTGGGTAATGAAGAGGTAATTTGGGACAGTGTGGTGGTAC
CAGGAAGAAAGAGGATTGGAAAGGCCAGTACTGTTTTAGTTGCTCGGCACTGTTGG
TTTTGTTTTAATGTGGTTGCCCTGTCCACTACATGGTTCTATCAGTAGTGTAATCC
ATTTTCAATGTAAAGCTCTTTTAGTTTTTGTCTATAGACATAAATTAATATTTTGAG
AGGCATCCCTCACCTGTTTCATTTCTTCTGTGTTGAAATGAAGTACTTAAATACC
GTTATACATGAACTTTGTGGACTGTAAGATTTGTTATATATGTTCAAATGCCTTTT
AGCTGGCTTTTTAATTAATATGCCTGTTTTGAGTGCTTAATACAATGTAATGNNGA
TTGTAAATCATACCTATTTTAAATCATTCCTTCCTGTATATTTGNACTCAGAGAGC
CTTATTTTATTCTTCCAGC

~~Sequence ID 491~~ SEQ ID NO: 134

nt:

382

TTTTCTTAGAACTTTATTTTTTCTGGCCAGGCGCAGTGGCTCACACCTGTAATCCC
AGCACTTTGGGAGGCCAAGGCAGGTGATCACCTGAGGTCAGGAGCTCAAGACCAG
CCTGGCCAACATGGTGAACCTGTCTCTACTAAAAATACAAAAATTAGCTGGGCG
TGGTGGCGCATGCCTGTAATCCCANCTACTCAGGAGGCTGAGGCAGGAGAATTGTT
TGAACCCGGGAGGCGGAGGTTGCANTGAGCCGAGATTGCGCCACTGCACTCCAGCC
TGGGCAACAGAGCGAACTCCATCTCAAAAAAAAAAAAAAAAAACAACCTTTATTT
TTTCTGATTTTAAAGTAATAACTAGTTTGTAGAAACATTAAAAGT

~~Sequence ID 492~~ SEQ ID NO: 135

ACCCTAAACATAACTTAAATTTGTTNGGAATTTGAAAGTACAGAATTTTCCTGTA
ATTGAGACTNTTTAACTTTTGTGGTTGGAGAAGGTATTCTATTTTTTGAAAATAT
CTGTAAGTTTTATCTAAATAGTAACTCTAAGTATTCTTCCCCTTTACTTACAGCC
ACCCTGGGAATCTGAGACTAGAGAAAATAAAGTTTGTCTCTTGTCTAAGGAGGGT
CTGGTTTAGAAATCTGATTTAGACATAGAAAAATTGCAAGAAGCTTGAGGTGATTG

Marked-Up Copy

GAAGATACGATTTTGTATCAAAGNATGTTTCTGTTTTATAGATTTTATTCATCTA
CAACTCCTTATTAATATATTTAAGAAGTCATTAACCCACCATTGATTACTTGATAT
AAAAGGAGAANC GGTGGTAAAAGGTGAAATANAATTTTAAATTTTTTTTTTTTAA
GTTTAGGATTTTTTTTTTAAATTCTAAGAGTTTCTGTCATTTGGGGACAATCAGAA

~~Sequence ID 493~~ SEQ ID NO: 136

TGGGAATCATAATTNGTTAACTGAAGCTNATAAGATGAGAGCATT CANAGAGAAAA
GAACGGAAAGATTGAATATCAGTTTCCCTTCTTTAAAAAAATGTGGATATGTGAT
CTAGCTTCTTGAGCATCACAGTGACTGATTGGCTCGTGGTAATTGATCGCTATGCT
GACAATCTTATCTCCACCTATGTCATTCAATTTTCTAAGAGGCAAAATCCTTAATC
AGGAGGAGAGTTTAGCTCTAGCTAAATTTCCCTTGTCCAGCATGCTCCTGCTCCCC
CAACTTGTGGAAACAGCTAAAGGATTGGACTAGGAGCANAAGTTTGGAATGGTTAA
AATGTAGCAACATGTGTTTCCTGAAACAAAATTCCACTATAATAAAAAAAGCATTT
GAATGCTCCCTTGTAATTCTGTTGGAGCTTGTTCCTTTTTTATGACACAACCATA
ATCAGTGATAGACAGTAGCATAAAGAAGCAAGAGCAAAGCAATTAAGTAATAATAG
CACTACAAAAATGTGTGCTGTACTTACCAAACACGACATTTATGAATTATTANATA
GGAATAAGGGGATGGT

~~Sequence ID 494~~ SEQ ID NO: 137

GACCCAGCCATCTAAATAAGTTRTACATGTTGCGTATTTTTTTGTTAGGGACTTAT
CTTCCGAAGAGGAAAGGTTTATGAAACCTAAAGTAACAATGATAGCTTGGAATCAA
AATGATAGCATTGTTGGCACAGCTGTGAATGATCATGTCCTCAAAGTGTGGAATTC
TTACACTGGACAACCTGCTTCATAACTTAATGGGACATGCTGATGAAGTATTTGTTC
TGGAGACACATCCCTTTGATTCCAGAATTATGTTATCTGCAGGACATGATGGCAGC
ATATTTATATGGGATATTACAAAAGGTACCAAGATGAAACATTATTTTAATATGGT
AAGTGAAGTGAGATGTACCTTGATACATGCTTGATAATTTGTTTAGAGTATTTGGG
TTATGCGGCTTACCCAGAAATTGATCTGCTTGTGTTTGGCAGTTTGTGTTTTACAAAT
CAACATATTCAAAGCCTGCTAAATATTAGACAGCTACATGTATATACGTACATACA
TGAA

~~Sequence ID 495~~ SEQ ID NO: 138

TTTC

~~Sequence ID 496~~ SEQ ID NO: 139

CTCGCTGGCGGGAGGCCACGGGCTTTCCACAGCGCGGGGAACGGGAGGCTGCAGG

Marked-Up Copy

ATGGTCAAGCTGACGGCGGAGCTGATCGAGCAGGCGGCGCAGTACACCAACGCGGT
GCGCGACCGGGAGCTGGACCTCCGGGGGTGATCTGGACCCTCTGGCATCTCTCAAA
TCGCTGACTTACCTAAGTATCCTAAGAAATCCGGTAACCAATAAGAAGCATTACAG
ATTGTATGTGATTTATAAAGTTCCGCAAGTCATAGTACTGGATTTCCAGAAAGTGA
AACTAAAATTTTAATCCAGGTGCTGGTTTGCCAACTGACAAAAAGAAAGGTGGGCC
ATCTCCAGGGGATGTAAAAGCAATCAAGAATGCCATAGCAAATGCTTNAACTCTGG
CTGAAGTGGANAGGCTGAANGGGTTGCTGCAGTCTGGTC

~~Sequence ID 497~~ SEQ ID NO: 140

GAAGACCTCACATCTGAGAGCTCATCTGCGTTGGCATTCTGGAGAACGCCCTTTTG
TTTGTAACTGGATGTACTGTGGTAAAAGATTTACTCGAAGTGATGAATTACAGAGG
CACAGAAGAACACATACAGGTGAGAAGAAATTTGTTTGTCCAGAATGTTCAAAACG
CTTTATGANAAAGTGACCACCTTGCCAAACATATTAACACACCAGAATAAAAAAG
GTATTCACTCTANCAGTACAGTGCTGGCATCTGTGGAAGCTGCGCGAGATGATACT
TTGATTACTGCAGGAGGAACAACGCTTATCCTTGCAAATATTCAACAAGGTTCTGT
TTCAGGGATAGGAACTGTTAATACTTCCGCCACCAGCAATCAAGATATCCTTACCA
ACACTGAAATACCTTTACAGCTTGTACAGTTTCTGGAAATGAGACAATGGGAGTA
AATATTACACAAATACTTATTCAATTGNGGTTATTTTTATACAGTAGTGAGAAGAAT
ATTGTTCCTAAGTTCTTAGATATCTTTTTTTGGATGTGCAAAAATTTTTTGATTGA
CAGTAACTTGGGTATACATGACACTGAAATGCCTTACTTTGGATGA

~~Sequence ID 499~~ SEQ ID NO: 141

TGCCTGCGGGCCAGGACCTCGCCAGCCCATGTTTCATCCAGTCAGCCAACCAGCCC
TCCGANGGGCAGGCCCCCAGGTGACCGGCGACTGAGGGCCTGAGCTGGCAAGGCC
AAGGACACCCAACACAATTTTTGCCATACAGCCCCAGGCAATGGGCACAGCCTTCC
TCCCCANAGGACCCGGCCGACCTCAGCGCCTCCTGCAGGCTAGGACACTGGTGCAC
TACACCCCATGCCTGGGGGCCGAGATTCTCCAGCAGAAAGATGCAATATTTTTTGT
TTCCTTTTTTTCCATTTTTTTCTCTAAGGAATCAATATTTCAATATGTTGAGTGTG
TGTCCAATGCTATGAAATTAAAATATTAAATAACATATTTATGGCATTTTCTTGAA
GAGTGTGGTTGAAGAAATATTTCTCCTTTTGTTTTTCTTTTTTTTTGNTTGNTAC
TGCCACTTCTTTTTAGGAGCAAATCTCCCCAGGGGTGTACGGNATTTCTTGACTCT
GGGAACAGCTGCTACCCCCAAGACTTGCCACGTTGTTCTGCCCTCAAATGGAATTA
AGTG

~~Sequence ID 500~~ SEQ ID NO: 142

nt:

Marked-Up Copy

GGAATATGGTCAGGATCTTCTCCATACTGTCTTCAAGAATGGCAAGGTGACAAAAA
GCTATTCATTTGATGAAATAAGAAAAAATGCACAGCTGAATATTGAACTGGAAGCA
GCACATCATTAGGCTTTATGACTGGGTGTGTGTTGTGTGTATGTAATACATAATGT
TTATTGTACANATGTGTGGGGTTTGTGTTTTATGATACATTACAGCCAAATTATTT
GTTGGTTNATGGACATACTGCCCTTTCATTTTTTTCTTTTCCAGTGTTTAGGTGAT
CTCAAATTAAGAAATGCATTTAACCATGTAAANATGANTGCTAAAGTCAGCTTTT
TAGGGCCCTTTGCCAATAGGTANTCATTCAATCTGGTATTGATCTTTTCACAAA

~~Sequence ID 502~~ SEQ ID NO: 143

ACCCGCCATCTTCCAGTAATTCGCCAAAATGACGAACACAAAGGGAAAGAGGAGAG
GCACCCGATATATGTTCTCTAGGCCTTTTANAAAACATGGAGTTGTTCCCTTGGCC
ACATATATGCGAATCTATAAGAAAGGTGATATTGTAGACATCAAGGGAATGGGTAC
TGTTCAAAAAGGAATGCCCCACAAGTGTTACCATGGCAAACTGGAAGAGTCTACA
ATGTTACCCAGCATGCTGTTGGCATTGTTGTAAACAAACAAGTTAAGGGCAAGATT
CTTGCCAAGAGAATTAATGTGCGTATTGAGCACATTAAGCACTCTAAGAGCCGAGA
TAGCTTCCTGAAACGTGTGAAGGAAAATGATCAGAAAAAGAAAGAAGCCAAAGAGA
AAGGTACCTGGGTTCAACTAAAGCGCCAGCCTGCTCCACCCAGAGAAGCACACTTT
GTGAGAACCAATGGGAAGGAGCCTGAGCTGCTGGAACCTATTCCCTATGAATTCAT
GGCATAATAGGTGTTAAAAAAAAAAAAATAAAGGACCTCTGGG

~~Sequence ID 503~~ SEQ ID NO: 144

nt:

109

ACATTTTCCGGNCCTTTTGCCATACACAGTTACAGAGATCAGTCAAATCCATACCA
CCACTGAGATCTCATTTATTGCCACAGATGCACAAAATAAATAACCCAAAATC

~~Sequence ID 504~~ SEQ ID NO: 145

nt:

374

CCAGCAACGACCCATACCTCAGACCCGACGGCCCGGAGCGGAGCGCGCCCTGCCCT
GGCGCAGCCAGAGCCGCCGGGTGCCCGCTGCAGTTTCTTGGGACATAGGAGCGCAA
AGAAGCTACAGCCTGGACTTACCACCACTAACTGCGAGAGAAGCTAAACGTGTTT
ATTTTCCCTTAAATTATTTTTGTAAATGGTAGCTTTTTCTACATCTTACTCCTGTTG
ATGCAGCTAAGGTACATTTGTAAAAAGAAAAAAACCAGACTTTTCANACAAACCC
TTTGTATTGTANATAAGAGGAAAAGACTGAGCATGCTCACTTTTTTATATTAATTT
TTACAGTATTTGTAAGAATAAAGCANCATTTGAAATCG

~~Sequence ID 505~~ SEQ ID NO: 146

GTACAGGAGGTAAATTGGATACCCCATCTAAGGGGATCTGTGAGACCAGGTAGTTA
TTTGAATGAAAGAGTAAGATATTAAACCAGCCAGCATGTCAACAGGTGGGTGATA
GTCTTGTTCTCACAGACAACAGATGGCCATCATCTTAAAACAACATTTATGTTAAC
CAGCAGATAAGGGACTCCTGCATTGTCAGTGGACTTTGAGCCTGAGTTTTTCTACT
TGCATAGGTGAAAGTGGACTGCAATGCTAGTATAAATGCCGTATGATGACTAGTAC
CCCTTAGGGAGCTCCAGTTTGCCCTTCTGGGGAACACAGACCCCAAGTGTAATTT
CCTGAGGACAGCCCGACTTCT

~~Sequence ID 506~~SEQ ID NO: 147

GTTACTGTGAGCCTGTCAGTAGTGGGTACCAATCTTTTGTGACATATTGTCATGCT
GAGGTGNGACACCTGCTGCACTCATCTGATGTAAAACCATCCCANAGCTGGCGAGA
GGATGGAGCTGGGTGGAACTGCTTTGCACTATCGTTTGCTTGGTGTTTGTTTTTA
ACGCACAACCTTGCTTGTACAGTAACTGTCTTCTGTACTATTTAACTGTAAAATGG
AATTTTGACTGATTTGTTACAATAATACTCTGAGATGTGTGAAAAAAAAAAAAA
AAAAAAAAAAAAA

~~Sequence ID 507~~SEQ ID NO: 148

nt:

521

CTGCGGTGGAGCCGCCACCAAAATGCAGATTTTCGTGAAAACCCTTACGGGGAAGA
CCATCACCTCGAGGTGAACCCTCGGATACGATAGAAAATGTAAAGGCCAAGATC
CAGGATAAGGAAGGAATTCCTCCTGATCAGCAGAGACTGATCTTTGCTGGCAAGCA
GCTGGAAGATGGACGTACTTTGTCTGACTACAATATTCAAAGGAGTCTACTCTTC
ATCTTGTTGTTGAGACTTCGTGGTGGTGCTAAGAAAAGGAAGAAGAAGTCTTACACC
ACTCCCAAGAAGAATAAGCACAAGAGAAAGAAGGTAAAGCTGGCTGTCCTGAAATA
TTATAAGGTGGATGAGAATGGCAAATTAGTCGCCTTCGTGAGAGTGCCCTTCTG
ATGAATGTGGTGCTGGGGTGTTTATGGCAAGTCACCTTGACAGACATTATTGTGGC
AAATGTTGTCTGACTTACTGTTTCAACAAACCAGAAGACAAGTAACTGTATGAGTT
AATAAAAGACATGAACT

~~Sequence ID 508~~SEQ ID NO: 149

AAGCTCATGATTTTAAATGTATTTTCTAATAAACTATACTCCCATTTAAAAATCA
CCAATACCTTAATGTTTCAATTATATAAGCTAATTAAAAATAAAGGCTGGGCGTGG
TGGCTCACTTTGGAAGACCGAGGCAGGCAGATCACCTGAGGTCAGGAGTTCGAGAC
CAGCCTGCCCAACATGGAGAAACCCATCTCTACTAAAAATACAAAATTAGCCAGG
CATGGTGGCACATGCCCGTAATCCCAGCTACTGGGGAAGCTGAGGCAGGAGAATCA

Marked-Up Copy

CTTGAACCTGGGAGGCAGGGGCTGCAGTGAGCCGAGATCATGCCATTGCACTCCAG
TCTGGGCAACAATAGTGGAATCCATCTCAAAAATAATAAAAAAATAAAATAAAA
ATAAAATTCAAACCTAAAATAGATGCTCTACTTCAGGAGTGGGCAAATTAATCACC
TGCATCCTTTTTTTTGGGCTTTC

~~Sequence ID 509~~ SEQ ID NO: 150

nt:

575

TTTTTTTCTAAATGGNGATTACTAATATATGTGGAGACTATTAATCTCTTTTCTGT
TGCCATTAGTTCATTTTTTCCCCAAAAGCCAATACATGTTCATTACAAAAATGAATT
ATAAAATATAAGTTAAAAGAAAAACATAAAACCCTACAATCTTACCCACCCAGACA
ACTACTATTAATACCTTAGTATTAACATATACACATCATGTATATGTATAAATTTA
TCTTAAACAAAAATAAAATTATTCTTTACATATTGTTTTAAACCTATTTATCTGG
CCAGGTGCCGTGGCTCACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCACGTG
GATCACCTGAGGTCAGGAATTCGAGACCAGCCCAGCCAACATGGTGAAACCCTGTC
TCTAATGGTTTTAAATACCAAAAAATTAGCTGGGCATGGTGGCACATGCCTGTAATA
TCAGCTAACATGGGAGGCTGAGGCAGGAGAATCACTTGAACCANGGAGGGGGAGGT
TGCAGTGAGCCGAAATCACACCACTTCACTGCAGCCTGGGCAACAAAGCAAGACTG
TCTCAAAAAGAAAAA

~~Sequence ID 510~~ SEQ ID NO: 151

CACGTGCATTCCCAGGAGGCTTTGGAGTCAGAACTGGATTCAAATTCTGACTNTAT
GTTGTGTGACTTGGGCCAATAGCTTCTTTNTGTGCCTCAGTTTCTTTAGCTGTAAA
TANACGGGTAGGTCACCCCTTACCCCATAGGTTATGGGGAAAGTTACAGAAAATGG
TCAGCTGGGCNCAGTGGCTCAAGCCTGTGGTCCCAGCNCCTTGGGAGGCCAAGGTG
AGCAGATTGCTTGAGCCCAGGAGTTTGACACCAGTNTGGCAACGTGACGAAACCCT
ATCNCTGTGAAAAATACAAAAAATTAGCCAGGCATGGTGGTGTGTGTCTGTGGTTC
CAGCTGCTTGAGAGTTTGAAGTGGGAGGATCACCTGAGCCCAGAAGGTCGAGGCTG
CAGTGAGCTGTGATCGCGTCACTGCACTCCAGCCTGGC-GACAGAGTGAGA-CCCC
T-TTTGAAAAAAAAAAAAAAAAAAT

~~Sequence ID 512~~ SEQ ID NO: 152

GTGAGCGGTGGTGGTTTATTCTTCCGTGGAGTTAAGGGCTCCGTGGACATCTCAGG
TCTTCAGGGTCTTCCATCTGGAATATATAAAGTTCAGAAAACATGTCTCGAAGAT
ATGACTCCAGGACCACTATATTTTCTCCAGAAGGTCGCTTATACCAAGTTGAATAT
GCCATGGAAGCTATTGGACATGCAGGCACCTGTTTGGGAATTTTAGCAAATGATGG
TGTTTTGCTTGCAGCAGAGAGACNCAACATCCACAAGCTTCTTGATGAAGTCTTTT

Marked-Up Copy

TTTCTGAAAAAATTTATAAACTCAATGAGGACATGGCTTGCAGTGTGGCAGGCATA
ACTTCTGATGCTAATGTTCTGACTAATGAACTAAGGCTCATTGCTCAAAGGTATTT
ATTACAGTATCAGGAGCCAATACCTTGTGAGCAGTTGGTTACAGCGCTGTGTGATA
TCAAACAAGCTTATACACAATTTGGAGGAAAACGTCCCTTTGGTGTTCATTGCTG
TACATTGGCTGGGATAAGCACTATGGCTTTCAGCTCTATCAGAGTGACCCTAGTGG
AAATTCGGGGGATGGGAAGGCCACATGCATTGGAAATAATANCGCTGCAGCTGTGT
CAATGTTGAAACAAG

~~Sequence ID 513~~SEQ ID NO: 153

TTTTTTTTTTATAAACTCCAATCATTTCCAGAGCTACTTAGCTCAGCATCTTTTTT
TTCCACGCTCTTAAGTTGTGTTTATACATTTTTTGATACAGTTAGATTGTTTTTGTG
ACATTCTTCATTCTATCCTGGGATCCCCCAACCACCTAAGTGGATTTTTTGATAAT
TTGCATGCTTTAAGGATAACTCTTCATTCTGNAAAGGGCTATGGGTTTTGGCAAAT
GCAGAGTCATGTATCCAAGATTACAATATCGCACAGAAGAGTTTCATCACTATATA
AAACTCACCAGTCTTCCTCCTATTCAACCATCTCCATGCCTTCTTCCCAGCCCTAA
CTCCTTAAAACCACTCATATCTTTACTATTGCTATAGTATTGCCTCTTCCACCATG
TCATATAAATGGAAACATACAGTATTAGTCTTCTCAAACCTAGTTTCTTTTACCTAA
CAACATGCATTTAAGATTCATAGTGTCTTTTAATGACTTGATAGATTATTTCTTTG
TAGCTGAATAAATATTGCATCTTATAGATGTAACCGTTTGTATATCCATATTTTCTC
ACAGCCTATGACTTGNCCTTTTGATTCTCTGAACAGGCCATTCACAAAGCAGAAGTT
TTAATTTTTATAAAGCTAATGNATCAACTT

~~Sequence ID 515~~SEQ ID NO: 154

CCTGGATGACAGCATATCTGTTTATAGCTCAGTTTACTGAATACTTTAAGCCCACT
GTTGAAACCTGCT

~~Sequence ID 518~~SEQ ID NO: 155

nt:

502

GATGCATGTCCAGCATAGGCAGGATTGCTCGGTGGTGAGAAGGTTAGGTCCGGCTC
AGACTGAATAAGAAGAGATAAAATTTGCCTTAAACTTACCTGGCAGTGGCTTTGC
TGCACGGTCTGAAACCACCTGTTCCCACCTCTTGACCGAAATTTCTTGTGACAC
AGAGAAGGGCAAAGGTCTGAGCCCAGAGTTGACGGAGGGAGTATTTAGGGTTTAC
TTCAGGGGCTCCCAAAGCGACAAGATCGTTAGGGAGAGAGGCCAGGGTGGGGACT
GGGAATTTAAGGAGAGCTGGGAACGGATCCCTTAGGTTTACAGGAAGCTTCTGTGCAA
GCTGCGAGGATGGCTTGGGCCGAAGGGTTGCTCTGCCCGCCGCGCTAGCTGTGAGC

Marked-Up Copy

TGAGCAAAGCCCTGGGCTCACAGCACCCCAAAGCCTGTGGCTTCAGTCCTGCGTC
TGCACCACACATTCAAAAGGATCGTTTTGTTTTGTTTTTAAAGAAAGGTGANAT

~~Sequence ID 519~~ SEQ ID NO: 156

CTGCGATNGAGTTTTGAGAGGAAGGANTAAAGTNCTCATCTCNGACGGTGAGAAAG
ATCATNACTAAGGAAACGCAGGGTTGGAAGCAGTGCTGANTGTCCAGTTGAGTTTC
ATGANCAAACATTTGCTGTGGGACCAGTTTTTCATGGNGGTTTGTCATTTTGTCCAG
CTGCCTGGAGCTGCTTGGTTGAAGGCACAGAATAATCAGGATTAATTGTTNAACTT
GTATGAATTTCTTTATTTTAAAATAGGAATAATATCTGCCTTGGGAGCAAGTTGTA
AGAGTTAACTGAAAGCTTNAGGAAAACTTCCCTTGCTATTTAAGTAGGGCTTTA
CAAGTTACAATTCTATCACAGTTTTTAAGATTATAAAC

~~Sequence ID 521~~ SEQ ID NO: 157

GCGGCGCANCTGCGGATCCANAAGGNCATAAACGANCNGAACCTGCCCAANNCGTG
TGATATCACCTTCTNAGATCCAGACNACCTCCTCAACTTCAAGCTGGTCATCTGTC
CTGATGAGGGCTTCNACAAGAGTGGAAGTTTGTCTCAAAAAA

~~Sequence ID 523~~ SEQ ID NO: 158

nt:

585

GATTTACTGTGGGAATTTGCTCATGCAATTATGGAAACCTAGAAGTCCCATAATAT
GCCATCTTCAAGCTGGAATCCCAGGAAAGCAGGTGGTGTAATTCTGAGATTGAAGT
CTTGAGAACCGGGGGAGTCAATGGTGTAACCTCCAATCTAGGGCTTAAGGCCCAAG
GACCAGGGCTGCTGGTGTGCAGATGCAAATCCTGGAGTTCAAAGGATTGAGAACCA
GGAGCTCTGGTGTCTGAGGGCAGTAGAAGATGGATGTTCCAGCTCAAGAAGGGAAA
GTAAGAATCCGTCCTTCCCTCCACTTTTTTGTCTATTTCAGATGAGCCCTCAATGGA
CTGAACGATGCTCACCCACACTGTGAGGGCTGGTCTTCTTTATTCAATCCACTGAC
TTAAGTGCTGATCTCTTCTGGAAACACCTTCACAGACACACCCAGAAATAATG TTC
TACCAGCCATGGGCCTGTTACTTAGCCCAGTCAAGTTGACACAGAAAATTAGCTAT
CACAACATCTGTGTGTGTATATACATATGTATTTGCATGTGTGTGTATATATGGNG
TATATATATTCATGTGTGTGTATAT

~~Sequence ID 524~~ SEQ ID NO: 159

CTTTTGCCAGTAGGCCCCCTGAGTAGGTTCTCTATCTTTTGGCATGACCCAGAA
GTCTTTGATAACTTCCTTGCTTTCTGATGTGACAAGACATCCAGGGCCAGATTGTC
CATATCCTGCCCCGGATGCACGATGCACTGTTTCTCCAAGAATCCCTGTGTCCTTT
GCTGATGATGCCATGATTTTAAGTTCTCTAATATAGTTTTATCTCTTTGTTTCAGA

Marked-Up Copy

TAATGCTTTTGTGTTCTCACATGTCCTGCTCTCTCTCTCTCTCATTTTGGTGTT
GATCAGTCTTTCCATAAGATTGTTTATTTCACTAGTCCTTCATTCTTCTTTTTTCT
AAATTTACTCTTCTTGACTAGTATCCTGTCACCTTCTGAGGACTCATATTTTGGCAA
CTTGAAAATTATTCTTATTTATTTAAGTATATGTTNCTGAAACTCTCATTAGACAC
ATTTTG

~~Sequence ID 525~~SEQ ID NO: 160

GTAAAAAAGTAAAAGGAACTCGGCAAATCTTACCCCGCCTGTTTACCAAAAACA
TCACCTGGTAGCATCACCAGTATTAGAGGCACCGCCTGCCCAGTGACACATGTTTA
ACGGCCGCGGTACCCTAACCGTGCAAAGGTAGCATAATCACTTGTTCTTAAATAG
GGACCTGTATGAATGGCTCCACNAGGGTTCANCTGTCTCTTACTTTTAAACCAGTGA
AATTGACCTGCCCCGTGAAGAGGCGGGCATAACACAGCTGAAAAAAAAAAAAAAAAA
AAA
AAAAAATTTT

~~Sequence ID 526~~SEQ ID NO: 161

nt:

516

CTTTTCATGGTCTCTTGTTTCATTAATCATCTAAAATCCAAGCNCAGAGAATTCAAT
TTTAGATGGTCTCCAGAGCAGAATTTGATGTATAATCTTAATTACAAATCATAGAT
AATTAATATTGNTTACAAAATCANAATACGATTAGAGGTAGGGATCCTGCACACAC
CCTATTTTCTCCCCAGTGTTCTGACCGAGAGACTAATTAATAATTCAAGGAACTT
ACAGTGAATGANAAACCATGGTTTTGCTTAATTATCAGAACAGCTAGATCTGAGAA
CAGCTGTCTCCCACATGGATAGACACTTATTCACCCATTTGCAGGTAGAATAGCT
GGCAATAATAAGTCCTTCCCATTGGATATGTTGAAAGGTGCCTGCCATGGCATAGT
TGCCACAAGAGAGGAAGAAATGGACACAAATGTAGGCTGTTTTCAGGGCANAGGGA
AGGTGGGAGGAAACCAANTTGCTGGTTTTTCACACACCCTCTGGGGAACACCCATGC
ACCTATGANATG

~~Sequence ID 527~~SEQ ID NO: 162

GACAAAAGCTGAGAGAATTTTTTTCTTGAATATTTGCACTAAAAGATAGGTAAAA
TTCTTCAGGCTGAAGAGAGCATAACCAGGTGGAGATTTGGATCTACAAAAGGAAGG
AAGATTTGGAAATGGATTTGGCACCATTGACTCAATTTCCAGAACAAGAAAGCAGG
GACAGTTTTGGGAAGCTCAAGACACACTGCCCATGAGCAGCAATTTGGACCTCTG
CTGCATCCACTGTGCATCAAACACACACTGTACAGACAAAGACTCCCAGGAAAAGA
AGTATAAACATGGACTAACACAGAGATGGGCAAACCTACAGCCTGTGACCCAGCCAC
CTGTTTATGTAGAATCCAAAGTAAGAATCTTTAACTTACACATAAACTT

~~Sequence 529~~ SEQ ID NO: 163; 660nt

GACAGCAGAGCACACAAGCTTNTAGGACAAGAGCCAGGAAGAAACCACCGGAAGGA
ACCATCTCACTGTGTGTAAACATGACTTCCAAGCTGGCCGTGGCTCTCTTGGCAGC
CTTCCTGATTTCTGCAGCTCTGTGTGAAGGTGCAGTTTTGCCAAGGAGTGCTAAAG
AACTTAGATGTCAGTGCATAAAGACATACTCCAAACCTTTCCACCCCAAATTTATC
AAAGAACTGAGAGTGATTGAGAGTGGACCACACTGCGCCAACACAGAAATTATTGT
AAAGCTTTCTGATGGAAGANAGCTCTGTCTGGACCCCAAGGAAAACCTGGGTGCANA
GGGTGTGGANAAGTTTTTTGAAGAGGGCTGAGAATTCATAAAAAAATTCATTCTCT
GTGGTATCCAAGAATCAGTGAAGATGCCAGTGAACTTCAAGCAAATCTACTTCAA
CACTTCATGTATTGTGTGGGTCTGTTGTAGGGTTGCCAGATGCAATACAAGATTCC
TGGTTAAATTTGAATTTTCAGTAAACAATGAATAGTTTTTCATTGTACCATGAAATA
TCCAGAACATACTTATATGTAAAGTATTATTTATTTGAATCTACAAAAACAACAA
ATAATTTTTAGATATAAGGATTTTCCTGGATATTGCACGGGAGA

~~Sequence ID 529~~ SEQ ID NO: 164

GACAGCAGAGCACACAAGCTTNTAGGACAAGAGCCAGGAAGAAACCACCGGAAGGA
ACCATCTCACTGTGTGTAAACATGACTTCCAAGCTGGCCGTGGCTCTCTTGGCAGC
CTTCCTGATTTCTGCAGCTCTGTGTGAAGGTGCAGTTTTGCCAAGGAGTGCTAAAG
AACTTAGATGTCAGTGCATAAAGACATACTCCAAACCTTTCCACCCCAAATTTATC
AAAGAACTGAGAGTGATTGAGAGTGGACCACACTGCGCCAACACAGAAATTATTGT
AAAGCTTTCTGATGGAAGANAGCTCTGTCTGGACCCCAAGGAAAACCTGGGTGCANA
GGGTGTGGANAAGTTTTTTGAAGAGGGCTGAGAATTCATAAAAAAATTCATTCTCT
GTGGTATCCAAGAATCAGTGAAGATGCCAGTGAACTTCAAGCAAATCTACTTCAA
CACTTCATGTATTGTGTGGGTCTGTTGTAGGGTTGCCAGATGCAATACAAGATTCC
TGGTTAAATTTGAATTTTCAGTAAACAATGAATAGTTTTTCATTGT

~~Sequence ID 530~~ SEQ ID NO: 165

nt:

660

GACAGCAGAGCACACAAGCTTNTAGGACAAGAGCCAGGAAGAAACCACCGGAAGGA
ACCATCTCACTGTGTGTAAACATGACTTCCAAGCTGGCCGTGGCTCTCTTGGCAGC
CTTCCTGATTTCTGCAGCTCTGTGTGAAGGTGCAGTTTTGCCAAGGAGTGCTAAAG
AACTTAGATGTCAGTGCATAAAGACATACTCCAAACCTTTCCACCCCAAATTTATC
AAAGAACTGAGAGTGATTGAGAGTGGACCACACTGCGCCAACACAGAAATTATTGT
AAAGCTTTCTGATGGAAGANAGCTCTGTCTGGACCCCAAGGAAAACCTGGGTGCANA
GGGTGTGGANAAGTTTTTTGAAGAGGGCTGAGAATTCATAAAAAAATTCATTCTCT

Marked-Up Copy

GTGGTATCCAAGAATCAGTGAAGATGCCAGTGAACTTCAAGCAAATCTACTTCAA
CACTTCATGTATTGTGTGGGTCTGTTGTAGGGTTGCCAGATGCAATACAAGATTCC
TGGTTAAATTTGAATTTTCAGTAAACAATGAATAGTTTTTCATTGTACCATGAAATA
TCCAGAACATACTTATATGTAAAGTATTATTTATTTGAATCTACAAAAACAACAA
ATAATTTTTAGATATAAGGATTTTCCTGGATATTGCACGGGAGA

~~Sequence ID 532~~SEQ ID NO: 166

GAATTGTGATAGTTCAGCTTGAATGTCTCTTAGAGGGTGGGCTTTTGTGATGAGG
GAGGGGAAACTTTTTTTTTTCTATAGACTTTTTTCANATAACATCTTCTGAGTCA
TAACCAGCCTGGCAGTATGATGGCCTANATGCAGAGAAAACAGCTCCTTGGTGAAT
TGATAAGTAAAGGCAGAAAAGATTATATGTCATACCTCCATTGGGGAATAAGCATA
ACCCTGAGATTCTTACTACTGATGAGAACATTATCTGCATATGCCAAAAAATTTTA
AGCAAATGAAAGCTACCAATTTAAAGTTACGGAATCTACCATTTTAAAGTTAATTG
CTTGTCAAGCTATAACCAACAAAAATAATGAATTGATGAGAAATACAATGAAGAGGC
AATGTCCATCTCAAAATACTGCTTTTACAAAAGCAGAATAAAAGCGAAAAGAAATG
AAAATGTTACACTACATTAATCCTGGAATAAAAGAAGCCGAAATAAATGAGAGATG
AGTTGGGATCAAGTGGGATTGANGANGCTGTGCTGTGT

~~Sequence ID 533~~SEQ ID NO: 167

CTTGAACCTCGGAGGCAGAGGTTGCAGTGAGCCGAGATCACGCCACTGCACTCCAG
CCTCGGGGACAGAGCAAGACTCCATCTCAAAACACACACACACACACACACACA
CACACACACACACAAAACAGATATACTGAACACAGCACAAAGTGGGACATAAGAG
ATTTAAAAGGGTTAGAGATGTAAATGGATCTAGGAATGGAAACCATAAGNGGGA
TTTATCAACTGGATTCTGCANAATGCTGTTAAGGCCAGATGTTAGCAGGTGTTACA
TAAAAAAGGGATACCATGAGCAAAAGTATTTGAACATGGGCAATGGTTGAAACAAG
TTTAAACAGATTATNTTTATTACCAAATCTCTCAAACCTTTAATATGCTATAAACA
TTGTGAAACAATAAAAAAACTTTCCAAAA

~~Sequence ID 534~~SEQ ID NO: 168

GGGAAGGGAGCTATGAGTGTGTGTGTTGTGTATGGACTCACTCCCAGGTTACCTG
GCCACAGGTGCACCCTTCCCACACCCTTTACATTCCCCAGAGCCAAGGGAGTTTAA
GTTTGCAGTTACAGGCCAGTTCTCCAGCTCTCCATCTTANAGAGACAGGTCACCTT
GCAGGCCTGCTTGCAGGAAATGAATCCAGCAGCCAACCTCGAATCCCCCTAGGGCTC
AGGCACTGAGGGCCTGGGGACAGTGGAGCATATGGGTGGGAGACAGATGGAGGGTA
CCCTATTTACAACCTGAGTCAGCCAAGCCACTGATGGGAATATACAGATTTAGGTGC
TAAACCGTTTATTTTCCACGGATGAGTCACAATCTGAAGAATCAAACCTCCATCCT

Marked-Up Copy

GAAAATCTATATGTTTCAAACCACTTGCCATCCTGTTAGATTGCCAGTTCCTGGG
ACCAGGCCTCANACTGTGAAAGTA

~~Sequence ID 560~~SEQ ID NO: 169

GGCGGAGGTTGCAGTGAGCTGAGATGGCGCCATTGCTCTCCCAGCCTGGGTGACAA
GAGCAAACTCCGTCTCAAAAAAAAAAAAAAAAAAAGCAATTTACTTAAAAAC
ATACAAACACAGAGACAAGTATTTTTGAGAAACAAATACCTTTTTTCATTTTTTATA
CCAATGTAACAATAATCCATTAAACACACCTTTACTAACTGTTTTCTAGGAGTCTG
ATATGATGAGGAAATAGGTAAACCTTTAATAGCCAGTACTAAATTAGAGTGGCACA
ACTTTCCTGTTGAGGAAAAAGATGGGTATTTTACTTTTCTGTTTTAGAAAAGTGGCTT
GACAACAGTATGCTTATGTCTTAGAGTTTGAAATTCAAGTTCTTGAACATTATTAA
TGGCTACAATCATTACATACCCACATTGGGCTGTATTCTTGATGAATCCAAAGTGAT
TTTCACCTCAACTCTGAATTTTCATTCTCCTCTTTTGAATATAATACAACCATCTCA
CTAGAGGAAGCATTTCAGTCTTTTCTGATTGGAGATTCATTATTGTTTTAGATAAT
GTTTTCATTTGCTTATGGGTATATAAAAAATTTTATCTTAAAAATATTTCTCTCA
TTTAGCTAGCAACATTGTTTTT

~~Sequence ID 561~~SEQ ID NO: 170

CTCAGGGTGATCTCTGAACCCAACTTGCCCCAAAGAAGGTTGCTCTGTCCTCTCC
ACATCCCCATCTCCTCCCTAGGGCCTTGTTGGGGAGAGGCTCCTCCATCTTTCCCA
AGTCACACCATCGTTTCTACGTGGTCTGGACAAGAGCAAGAGCACACCTTGTTCCC
CACCTTCTCCAGAGCAGCCAGAACCCACCTCAGGTGCCTTCCCCATCCGGTGCACT
TAAGGCACTTCTGCCAGCACCATGGTATGAGCACTAGACTTGAGGTTAAGATTTGA
GAGCCCCCTCTGTCACTGTGGAAGCTTGAGCATGTTGCTTGATCTCTCTGAACCTT
GTGTTTCTCATCTGTGAAAGGTGATAATGTGGGGCTGCTGTGAGATTTAAAGGACA
TAATGCACCTACGGTCCAAGCACTGCCTGGAATACAGCANAAAGCTCAACAGATACT
GGACAACCCATCCCCTTAGTAGAGGCACTAACCATGTGACCCAAGGCAAAAGTGCT
TAAAAAAA

~~Sequence ID 562~~SEQ ID NO: 171

nt:

580

ATTGCATGCAAGTTTGCTGAGCTGAAGGAAAAGATTGATCGCCGTTCTGGTAAAAA
GCTGGAAGATGGCCCTAAATTCTTGAAGTCTGGTGATGCTGCCATTGTTGATATGG
TTCCTGGCAAGCCCATGTGTGTTGAGAGCTTCTCAGACTATCCACCTTTGGGTGCG
TTTGCTGTTTCGTGATATGAGACAGACAGTTGCGGTGGGTGTCATCAAAGCACTGGA
CAAGAAGGCTGCTGGAGCTGGCAAGGTCACCAAGTCTGCCCAGAAAGCTCAGAAGG

Marked-Up Copy

CTAAATGAATATTATCCCTAATACCTGCCACCCCACTCTTAATCAGTGGTGGAAGA
ACGGTCTCAGAAGTGTGTTTCAATTGGCCATTTAAGTTTAGTAGTAAAAGACTG
GTTAATGATAACAATGCATCGTAAAACCTTCAGAAGGAAAGGAGAATGTTTTGTGG
ACCACTTTGGTTTTCTTTTTTGGCGTGTGGCAGTTTTTAAGTTATTAGTTTTTAAAT
CAGTACTTTTTAATGGAAACAACCTTGACCAAAAATTTGTCACAGAATTTTGAGACC
CATTAAAAAAGTTAAATGAG

~~Sequence ID 563~~ SEQ ID NO: 172

GCAACCTGCACAACCCCGCCCTGTTTCGAGGGCCGGAGCCCTGCCGTGTGGGAGCTG
GCCGAGGAGTATCTGGACATCGTGCAGGAGCACCCCTGCCCCCTGTCCTACGTCCG
GGCCACCTCTTCAAGCTGTGGCACCACACGCTGCAGGTGCACCAGGAGCTGCGAG
AGGAGCTGGCCAAGGTGAANACCCTGGAGGGCATCGCTGCTGTGAGCCAGGAGCTG
AAGCTGCGGTGTCAGGAGGAGATATCCAGGCAGGAGGGAGCGAAGCCCACCGGCGA
CTTGCCCTTCCACTGGATCTGCCAGCCCTACATCCGGCCGGGGCCCAGGGAGGGGA
GCAAGGAGAAGGCAGGTGCGCGCAGCAAGCGGGCCCTGGAGGAAGAGGAGGGTGGC
ACGGAGGTCCTGTCCAAGAACAAGCAAAAGAAGCAGCTGAGGAACCCCCACAAGAC
CTTCGACCCCTCTCTGAACCAAAATATGCAAAGTGTGACCAGTGTGGAAACCCAAA
GGGCAACAGATGTGTGTTTCAGCCTGTGCCGCGGNTTG

~~Sequence ID 564~~ SEQ ID NO: 173

nt:

671

GGAATAGAATTTTAAATAGTAATAACTGCTTGTTTTTTTTGTGCAAGTACTTTTAT
ACATAAGATAAAACAAAACCTTACCACCAAACATACCAAAATGCACCTCTTTCATA
AGTGAGTTACTAAGATTTCTATACCTGGAATATCATGTATGTTTCATTTACTGGAT
GTTTACATTTTtagGAAGGAAAATAGTTTTGTTTTATTTAAACAACCTGAATACTTATA
AACTGTTGTTCCCTGGAAGTTATTTATTCATAAAAAATTTGTTCTTTTGTGTCATGAA
TTTATAATTCCTAAATGAAGACCAGAAAGTACAAATTGCTGGGAGGAAGAATAGGC
TTTATTAATCAACTGATGTCTTGATTTTTCTAAATGGGAAGATTGCTTTATTTTTTA
ACACTAATTATGGGAGCAGATTCTTAGCAAACCTCTTTGGAAAAGTTAATGTTATG
ATGTGCATTAGGCTGCCCCATCGTGTATATAAATGAAGCAGATTTGATTTTTGTAT
TCTTACGTTTCTCTGCTTTGTAGTTGTGGCTGTACTTAAAGAAATACAGAATTTCA
TATATTTAAAAATGTTTAAAATGTGACCCACAGACATTGTAAATGGATTNAAAAC
AACATGAAAAATATTCAACCTAAAAGAATTCTTAACTTCACAAGTGTTTTACTTC

~~Sequence ID 565~~ SEQ ID NO: 174

CTTGTTCCGCGTTCCTGCACAAAATGCCCGGCGAAGCCACAGAAACCGTCCCTG

Marked-Up Copy

CTACAGAGCAGGAGTTGCCGCAGCCCCAGGCTGAGACAGGGTCTGGAACAGAATCT
GACAGTGATGAATCAGTACCAGAGCTTGAAGAACAGGATTCCACCCAGGCAACCAC
ACAACAAGCCCAGCTGGCGGCAGCAGCTGAAATCGATGAAGAACCAGTCAGTAAAG
CAAAACAGAGTCGGAGTGAAAAGAAGGCACGGAAGGCTATGTCCAAACTGGGTCTT
CGGCAGGTTACAGGAGTTACTAGAGTCACTATCCGGAAATCTAAGAATATCCTCTT
TGTCATCACAAAACCAGATGTCTACAAGAGCCCTGCTTCAGATACTTACATAGTTT
TTGGGGAAGCCAAGATCGAAGATTTATCCCAGCAAGCACAACTAGCAGCTGCTGAG
AAATTCAAAGTTCAAGGTGAAGCTGTCTCAAACATTCAAGAAAACACACAGACTCC
AACTGTACAAGAGGAGAGTGAAAGAGGAAGAGGTTCGATGAAACAGGTGTAGAAGTTA
AGGACATAGAATTTGGTCATTGTTCACAAAGCAAATGTGTCGAGAGCA

~~Sequence ID 566~~SEQ ID NO: 175

GTCACCAAGAGCTTGTTGTCAGGTTTTCACTTGCTATTCGCAGAGATTTTTTTTAA
AGGCACTATTTGTAGTGTTAAAAGGGTGAATTTATCANAAGGCATAATAATCATAA
ATGTGTATATGCCTAATAATAGAACTTTAAAAGGCATGAAGCAACACTCAAAAGGA
TTAAAGGGAGATCATCTCACCCCCTTCTTACCAATTGATAGAATGATCTGATGAAA
ACAGTAAAATAACAACAGATCTGAACACTGTCAACCATCTTGACAAATACTTATGC
CTAGTGTTCCATTATTGGAACACTAAACATGTGGAATGATTTATATCCTACTGCTC
AAGGTCATCACCAAGGTCTAATTGTAAAATTTCAAAAATTGCAACCTCAGGCATA
AATGGGTTAATCGACATTTATAGCACACACATGCAACATGTACCAGAGATTCCTTC
TTTTCTATGAACATGGTACTTCCACCAAGATAGACCACATTGTGAACTATAAAACA
AATCTAAAAACATTTGAAATGAAGGAAATTATATAAAATATGTTCTCTTGATCTCA
ATGAAATTAAATTAATACTATAT

~~Sequence ID 567~~SEQ ID NO: 176

CTCATGGCGGCCAATGTAGGCCCAAACTTCCTCAAGTCAAACCTCTCCAGGCCCAC
CTTCTGCTTCCCGGTGGCATCAACAGGCCCAGCTTTGACTTGAGAACAGCCTCTGC
AGGCCCTGCTCTTGCCCTCCCAGGGGCTTTTTCCAGGCCCAGCTCTTGCCCTCATGGC
AGCTGCCCCAGGCCAAATTTCTGCCTGCCTGCCAGCAGCCTCAACAGGCACAGCTC
CTCCCTCACAGTGGCCCATTTAGGCCCAACTCATGACTGTGAGGCCATTTCCAGGC
CTAGTGCCTGCCTCGTGGCTGACTCTTGAAGCCCCAAACTTCCTCAAATCAGCCTT
TTGCCCAACTTCTGTCTACTGTTCGGACTCTACAGGTCAGCCTCTGCCTCACAGTGG
ACCCTCCAGACCCAGATGGTGTCTNCTGTGGCATCCTCAGGCGAAGCTCCTGCCTT
TCGGCAGCCTCTCCAGGCCCAGCTCCTCCTGCTCCAGCCTTCTCTCCAGGCTCTGA
ACTTTCTCAGGTCTCCCTCTGTTGTCCAAGGCTGGAGTGTAGTAG

~~Sequence ID 568~~ SEQ ID NO: 177

TATATATGTAATGCCCTTAACCTAGTGTTTGGCATGATCGTTGCTGAAAGGGAAGC
TTGTGGGTACAGTGTCCCCTCAGAAAGCCAAAGCCCAGGGAAGGTCGCCTGCCCAGG
TCAGGCTCCCAGCGAGTTTGTCTGGGGAGGGGGCCATTTCATACCTCCAGGTCAGGAC
AGAGGCTCGGGCTGAGGGAACCCTACACAGGTCCTGGAAGCAGATCCTTCCTGCCT
AAGCCAGCAGGACAGCTCAACAGGAAGCATCTTCCAGCCACGGGAGGAGAGGCAGC
ACCTTTTTTTGGAACCATAACAGAGCTAAGAATGGTGGTACAAGTAATAGATTCTGTA
CTGGCAACCCCACTTGGTGGAGCAAGTTCTAGGAAAAGGGGGCTGTCCTTGAGTCA
GCCATGGGGTCAGCCACACAGTCACCGCAGCTGCTCTTTGGCACCGGGCGCTGGAA
AGACCTAGGATGACACAGCCTGGAAAGAGCTTGGGAAAAGCTCATCTTCCACAGAA
CTACCTGCTATACCAGCCAGGGCAGGTGCTTATTTCCACAACAGCCCTCTGTTGTA
GGCGGCAGTGCCATCCTGAANGTGCCGTGGTACCTTCTGAANACCCAGCTGAGGGC
CTGTAATGGCACTTGCATGCCACATGGNACACCCTTTCCCGGTAA

~~Sequence ID 570~~ SEQ ID NO: 178

ACCGCGGCCGCGTNAANAAAAAAAAAAAAAAAAAGAAATCCACTTGATCAACTTAATTCC
TTNTCTTTATCTTCCCTCCCTCACTTCCCTTTTCTCCCACCCTCTTTTCCAAGCTG
TTTCGCTTTGCAATATATTACTGGTAATGAGTTGCAGGATAATGCAGTCATAACTT
GTTTTCTCCTAAGTATTTGAGTTCAAACTCCTGTATCTAAAGAAATACGGTTGGG
GTCATTAATAAAGAAAATCTTTCTATCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA
AAA

~~Sequence ID 571~~ SEQ ID NO: 179

nt:

457

TTAGAGAGGTGAGGATCTGGTATTTCTGGACTAAATTCCCCTTGGGGAAGACGAA
GGGATGCTGCAGTTCCAAAAGAGAAGGACTCTTCCAGAGTCATCTACCTGAGTCCC
AAAGCTCCCTGTCCTGAAAGCCACAGACAATATGGTCCCAAATGACTGACTGCACC
TTCTGTGCCTCAGCCGTTCTTGACATCAAGAATCTTCTGTTCCACATCCACACAGC
CAATACAATTAGTCAAACCACTGTTATTAACAGATGTAGCAACATGAGAAACGCTT
ATGTTACAGGTTACATGAGAGCAATCATGTAAGTCTATATGACTTCAGAAATGTTA
AAATAGACTAACCTCTAACAACAAATTAAGAGTGATTGTTTCAAGGTGATGCAATT
ATTGATGACCTATTTTATTTTCTATAATGATCATATATTACCTTTGTAATAAAAC
ATTTTTC

~~Sequence ID 572~~ SEQ ID NO: 180

CGTCTATTTGNGTTTCTTCTCACAATTGGTAAGTTCTCTGTATTGATTGATGGCTA
AGTTTGATTAGTGTTTTTCTCTAGTTGGTAATTATATTCTAGTATTTTATCATCTT
ATTGTTTACTCAACTNAAAGTGNACAGAAGAGTTGCCAGGTTTCTCTTTGATATG
AGATCTCTNNTTGATTTGGAATGCAAATCANAAGTGTCATGTTTTGAATAAAGGGA
CCAGATGACTTATAGGTATTCTTTCTCTAAATATAACTAAGGTAAGATTTTTGTTT
TGAGGTACTTAATCTATATAAGTGGTAAAGAATTTACTTGAATTTCTCCAAATTCT
CATGTCTAAAGTCTGATTGATTAAATTCATTCTTGGTATTTTCAATTTGAAAAGAAT
GTAGCTTTAGCAAACCTCTTTGTATAAATGCAGTGGGATTAAGGTCATTTAAAAAA
TTGTTATATCATTTGTATTTTTTAAAAATTTACCAGTTTTATTTTTCTTTTTACCCTT
AGCCCGGCCTCAGAAAGTGTGTTTGTGTCCATTTCTCCCAGCGCACCTCTGCATA
TCTCTACCCACTTGTCTATAATTCAGCATCCAGCAGAGGAAAACAAAGTGTTGCGTA
CAGTTCCTCTACTAGCAGCATGCCTCCCCCAGGACAAGTGTA

~~Sequence ID - 574~~ SEQ ID NO: 181

TTATTGCTGACATAAAAATGGTGCACATCGGCCAGGGCCCAGGATGAATCAGCCAA
TCTGCACCATTTATACATGGAACCTGGAGAACATTGTGCCAATAATCATTTAATATA
TGCCAAATCTTACACGTCTACTCTAAACTGCTCTAATGAAGTTTCAGTGACCTTGA
GGGCTAAAGATTGTTCTTCTGGGTAAGAGCTCTTGGGCTGGTTTTTCANAGCAGAG
TTCTTGTTGTGGGTAGACTGTGACTAGGTTTACAGCCTTTGTGGAACATTCCGTAT
AACGGCATTTGTGGAAGCAATAACTAGTTCCTATGAAAGAACCAGAGCTGGGAAGAT
GGCTGGGAAGCCAGGCCAAAGTGGGGGCAACAGCTTGCTTCTCTTTCTCTTCTCAC
CCTCAGTTTGTATGGGAAAATGGAGATGTCCTCTCCACTTTATCCCACGATATCTA
AATG

~~Sequence ID - 575~~ SEQ ID NO: 182

nt:

209

CAGGATATCGAGACCATCCCAGACAGCATGGTGAAACTCCGTCTCTACTGGAATAC
AAAAAGTTAGCCGTGTGTGGTGGCACGCGCCTCTAATCCCAGCTATTCGGGAGGCT
TAGGCAGGAGAATTACTTGAACCCGGGAGGCGAAGGTTGCAGTGAGCTGAGATCGC
ACCATTGCACTCCACCCTGG-CGACAGAGCAAGACTCCGTCT

~~Sequence ID - 576~~ SEQ ID NO: 183

nt:

541

CAGCCAACCCAGAAGGAGCCAGTCTACAACTATGCCTGATCCTCCTCATGGCAGGC
CACGAAGCATTGCTGCCATGTGTTGAATTATAAAACCCACATTGCTTTTTGAACCC

Marked-Up Copy

TGTTGCGGGTAAAAATAACCAAATTATCAGTCCTTGGAAACCCAGGCAATCAAGTG
AGTACAAGGTAAAGATAAGTATGGTTTAGAGGAGAAATTATGTTCCCTGAACTGGTG
TCCTTTGATGGCAGCGTCAGCCTTGCTAAGTCAGAGTAGAGGGAGCAGTGACCTTA
ATAAGCTTTGGTGAGCATCATGTGCACGCGTGGGTGGGAGTCCCTTTCACTGATGC
TTTTAAAAGTGCTTTTGCAGACCCTGGAAGGGATCCTCCACACATATGAGGTGTGG
GACAGGTAGGCCAGAGAGGATTAGCCCTGCTTTTCGAGACTAGAAATCTACAGTCCT
GAAGGAGCAGTAATTAATTGGTACACCTGTCAGGGCCAGCCCCCAGGTCTCCTGGC
TTTTTCCAGGTTTTCTGTCTCACATGATTTTGCTTTT

~~Sequence ID 577~~SEQ ID NO: 184

CTTTAATTTTTCAAGTGTTTAAAAACAATTTTATACTTAAGCCAGCCTTGAAGAT
AAGCACAAAATTTACCAGTTTACATTTAAAAACAACAAAAACGACAACAACTC
AAGCACCCGCTCTGTGCATAGCACTATTCTAGGTGCAATAAAAGGGAATCTTAACC
TTAGAAATATGAGTTCACCTTTCTGGAATTGTATTATCTCCTTTTCCAGAGAGTAAA
AATAAATAAAATCACCATTGTTTACTACAGATCTGCCCCAAACCACATCTGGTTCA
CAGAAAGGCTAATTTCTGCCAAATTAAAGATGTAATGAACTCAGTTCCTGCTTTCC
CAAAAACACGAAAGCAGAATTCCTTTTCACTGAAAAAATAAACAGTTTTCCATGC
AAGGGCAGTTTGCTTCTAATAAGTATTTTTTAAAAAATTTTTTTTTCTCCTCTAGCTT
TTCTTTAAATTTTCTCCTCTAATATTGCCTTTTCTTGTACAAGGCAGACCAGGTA
TCTTTTTATGCTGTTTTTCTTTTACTAAGAAAAGTATTGCATCTTGAAGACAAACC
ATTTCCCAGAGTAGTGATAAAAAATAACACTAAAAAACTTTAAAGGTGAGTCACT
TCATCACCTTGATGAAGTAAAAAA

~~Sequence ID 578~~SEQ ID NO: 185

GGAAAAAATATTTCCACTTAGATATTTTACATGGTTTTGTTTAAATTACCATTAC
TTGTTTTTTAAAAACACATGACCACATATGTATATGTATATCTACCTAAACATTGT
ATCATGGTTTCAGTATGTTATTCATGTATTACTGGGAGATGCTACCAAGAAACCAA
CCCAAAGAAAATTCTGGAAAATACATTTCTATTTATAGAATAAATGTTTCATTTAT
ATAAAAGCAAAAGAACTTAGAGTTCTAATAAATGGGATGTCTAATAAATTATGAAG
TTACTGATTTGAATATATTATATTTTTTATAACTTCCTTGCCAAAGTCCTGATTTAG
TACATTAGAGAACCTGTGTTTCCTCTCTCCTCTACCATTCATCTCTCTTCCATACA
GTCATTTGGGCTTTTTTACTCAAAGAGAATCAAGAAATAATAAGGTATAACAAGCTT
GGCAAAGTGTTGGCTTTTTTAAAAAATAATTTTTTAAATCTCTAGCAGTTTGGTAAT
TTAGCAGCATCATTTATTTGGGATTCTTTTATCTGATTTCAACAGTGAAAAACATC
CCTATGATAAAGCCTAATGACCCATTTCCAAAAGATGGAATTGCCCTTCCTAGAAA
ATATGACGGAGAAAAGT

~~Sequence ID - 579~~ SEQ ID NO: 186

nt:

502

CGAATAGCCAAGTGGTCTGACAAGATCGAGAGTAATGAGGCCCATACTTTAGTACA
GTCTTGAATGGCCAGATGGTGCTGGGCATACCCCAACCAGAGATATGTAAGTCTTT
ATGTTGTCAAAATTTCCCAGAAACATGAATTTCCCCTAAGATTTCATTANAAAGAAATGAAGAA
CTAGAATGAAAACAAAAACGTTTCCTTGTATAATATTCATTANAAAGAAATGAAGAA
GGCCGGGCATGGTGGCTCACGCCTGTAATCCCAGCACTTTGAGAGGCCAAGGTAGG
CAGATCATGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATAGTGAAATCCCGTC
TCTACCAAAAATACAAAAAATTAGCCGGGCATGGTGGCACACACCTGTCATCCCA
GCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGTGGAGGTTGCA
GTGAGCTGAGATTGCACCACTGTACTACAGCCTAGGTGACAGTGCAAGACTCTG

~~Sequence ID - 580~~ SEQ ID NO: 187

nt:

316

CCTATGCCAACTAAAGAAAGCTTGCCTGGCCTACAGGCCTAAAGGTTCAAATGNG
GATTAAAAAACACAGTAGTCACATAAAATGTCTGCTGGCTGGCTGGAATTCCATC
ACCTACAATTTACCTGCTTTCAAAAACCTGTGTTCAACATTGAGAAAACAGAAAACC
ACTTATCTTGAGCTTAATATGGGCTTCTTTTTCTTAACTGTAGAACACTTACTGA
AATATCAAATCAATGGTTAGGATATGTATCCTAGGCAGGCCTAAACCATTAACACT
TGGTTTAAAGCAACTTTGTATAATTNACCTCCTAAAT

~~Sequence ID - 581~~ SEQ ID NO: 188

CTTCATGAGTGCCCGGTTGCCCAAGTCAAAAACCTGGGAGTGATATAAACTCCCCA
CACATCCAGTCAGTCACTCATCAACTCTATTGATTCTG-CTGCTAAATATATCTCA
ATTGTATTAACTTAAACATATGCATAATACATCTTCTTCTTCACTGCATTTTTGTG
GGCTGCACTTACCTTTTCAAGTAACAACAACACTGGCCCCCTCTTGCCCTTCTAGTCA
GAAGTGCCAAAATGATGAGAGCTAGCCATGACAAACCCACAGCCAACATTACACTG
AATGTGCAAACTGGAAGGGCATCCAACAGAGGAGG

~~Sequence ID - 582~~ SEQ ID NO: 189

TAGAATTCTCGCCTGCCTTGGCTTCTCCCTCTAGTTGTTTCCTTCTCTGTCTTCTGT
GGGCTTCTTATTGTCTGCTCACTCCTTCTTTCAGTGTCTCTCATGGGCTTCCTTCC
CTTCTCAGCTGATGCCATCACCTGGGGAATCACAGTTACTCAGCAGCACTGGGGCC
TCTCTATCTCTATGCTGGTCATGCCTATGTGTGAGCTGCAGACCCAGTGGAATTTT
CATTTGTGCATCCCATGCCAGCCCACCCTCCACCAGCCTCGAATGCAGCTGTTCA

Marked-Up Copy

GCCCTACCCCAGTCCTCAGAAAAGTTCCCTCTCCCTGGATCCTCTTTTTTCCTTCATG
AGTGCCCGGTTGCCCAAGTCAAAAACCTGGGAGTGATATAAACTCCCCACACATCC
AGTCAGTCACTCATCAACTCTATTGATTCTGTCTGCTAAATATATCTCAATTGTAT
TAACTTAAACATATGCATAATACATCTTCTTCTTCACTGCATTTTTGTGGGCTGCA
CTTACCTTTCAGGTAACAACAACACTGGCCCCTCTTGCCCTTCTAGTCAGAAGTGC
CAAAATGATGAGAGCTAGCCATGACAAACCCACAGCCAACATTACACTGAATGTGC
AAAACCTGGAAGGGCATCCAAACAGAGGA

~~Sequence ID - 583~~ SEQ ID NO: 190

nt:

631

CTGAGGTGGGAGGATTCCACTCTCACCCATTTCTTCTTTCATTTTCAGTTTCTCCA
GTTAGTAACTGAAGATGTTCTTTGAGTAATTAAGTGAGTGAGAAAATTTTAAAGTG
AGAAATCTATAAAAAGAACCATGTTAACATAAATATTTTCAGTCCTTACAAGTTGGT
ATTGACTTTTCTCATTGGTAATCTGACTGATTTAATACTGCTCATTCCAATATCTG
GTGATGTAATTCTGGTTATGAATCCTTGTATTAATAACACCTCCTGGGAGGTTTTT
TTTCCCCAACATTACATTCAGAATATTAGAGCTGAAAATACCTTTTTTTAAGGTTAT
CAGGAGGAGGGAGCTTATGTTTAATGTGGTGGATAAACTTAACTGCTGGTTAATA
CAATTGTTATTTCAGGTGAAATTCCCTAAACTTTTCACGTGCAAAGTTTTGTATGTA
TACAGACATTTGGGGAAAAGTTTTATCATCCCTAAAACCGGTTACTGTCCAGAAAA
TGATAAGAATCCCTGGGTTCCAAATCCTTCATAAGGTATTTATTCATTTATTTATT
CAACACATTTACTCAATGCCTCCGCTCTGCTGCAACTACACTGACATTCTGCTTCT
AATCTAACCGAAAAT

~~Sequence ID - 585~~ SEQ ID NO: 191

TTTCAAATTGTACAATAACACAAACAACTTTGTTAAGGCCATGTTTTATTTGCTGA
TTAATGGACAAAAGGCAATGTAATTTATTTTCAAGTATTTTCTTGAAAGTCTGTGC
TCATAAAAATCATGAAAAGTTGGAAAGACTGTTAAATCACTGAACTTCAAATATA
TCTTACACAATCTTGTTTGTACAAAAATACAAGTTAAATATAAACATAAAGCAATC
ATGGTAATTTTATGCAAATCTGTTTTATGTGATCATCAGTTATATATAAAAGTTTC
TCAGTTCTGTTATTTGTGAAAAGATCAATACCAGATTGAATGACTACCTATTGGCA
AAGGGCCCTAAAAGCTTACTTTAGCACTCATCTTTTACATGGTTAAATGCATTTTC
CTAATTTGAGATCACCTAAACACTGGAAAAGAAAAAAATGAAAGGGCAGTATGTC
CATAAACCAACAAATAATTTGGCTGTAATGTATCATAAAACACAAACCCACACAT
CTGTACAATAAACATTATGTATTACATACACACAACACACACCCAGTCATAAAGCC
TAATGATGTGCTGCTTCCAGTTCAATATTCAGCTGTGCATTTTTTCTTATTTTCATC
AAATGAATAGCTTTTTTGTCACC

~~Sequence ID 586~~SEQ ID NO: 192

GTAAACTGTTCTCTCCGAGGGAAAAAATGGAAGTTATCCTCACAGTTCACTGCCGT
GGTATTTCTTCTGTCCCATGCTTTGCATGACTGCCATGGTACAGCCTTGTTTCAA
CTGTTCACGTGTGATCTGTGGGTCTTTGAGTTTCAGTGAGTTTGCTGAAATGTCGAA
GAAGTAGTTCCAACTTCAATGTTCAATGAAATTTTTGTTCAAGTTTGAAATGGAG
AGAGCAGCTTTAAAAGGTACTAAGCCTTTTACAAATTGGTGAGTACTGGCACATGA
GAT

~~Sequence ID 587~~SEQ ID NO: 193

TTTTTTTTTTTTCCTTAAAAGGTAACCCCTAAACACAGCTAAAACTATGCCATCAGC
TGACTCCAAGGNACACACAGTCCTGTATCTGGAACACTGAGTGGCAGGCATCTTT
CTCTGCCTCTGACAGTGGAGTCCCCATCACTGCAGAGCATAGCCAAAGGAGTCAA
GGTCTCAGCGGGTCACTGCCTTATCAACCCCTACCCAGTCCCTTATGTTTTTTAATA
TTTTATAATCTTGACATGACACCAAGATGCTTTAATAAAAAAGCACCTCTAACTCG
GTCTTGTTACTTACCTTGAGCCTGGGACTTCTCTAGGCTCCTGAGGCAAAAAC
AGGTAGAGGGGAGATGGTGGAACATAAAACACAATTTTGCTTGGCACCCACCTTGG
CGTCTGTCCCATGACCAGGTCTTTCAATTCGATGATTTTGTCATTGATGGAGGAG
CGATATCGTTTCTCAATGATATTATGGGTGTCCGCCTTTCTCCTTCTTTGGGGGG
CTCAAGCTGCTTGACTCCCCCAGGTACCTGCTTAATGGGGCACTTTCTCTTGCCCC
ATCATTACAGGCATTGTGGTCAGAAATGGTCCCACTGCTGCCCAACAGGGTCTA

~~Sequence ID 588~~SEQ ID NO: 194

CTAGTCTTTTCATAGTCTGCATAGAGTCTGGCCATTACCATCAGTTTTTAAGATGT
CCATATTGTGGCCGGGCGCGGTGGCTCACGCCTGGTAGTCCCAGCACTTTGGGAGG
CTGAGGCAGGTGGATCATGAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTG
AAACCCGTCTCTACTAAAAAAATATTAATAAAATTTGGCCAGGCCTGGTGGTGGGCG
CCTGTGGTCCCGGCTGCTTGGGAGGCTGAGGCAGGANAATGGTGTGAACCCGGAAG
TCGGAGGTTGCAGTGAGCCAAGATTGCACCTGGGCAACACAGCGAGACTCCGTCTC
AAAAAAAAAAAAAA

~~Sequence ID 589~~SEQ ID NO: 195

CAATTATTTATTACCTTTCCATTTGTTTCGCCTGATGATGTGACAATGCATGGTCTT
TGTGCATGCTGCTAGACACTTTTCTTTCCCAGCCGAAAAGTCTATTATGTAATTTT
TACATTACATAATTTTAATGTGGATGATCAGGATTAAATCAAGATATATATCTGGAA
CCTCTTATAAATGGAGCACTTAGAAATTTGTTGTTCTGCACTTAACCTAGAGAGAG

Marked-Up Copy

AAAAAATGCTTTTCTTTGTGAAAAATCTGAATTCCTGTCCTGACCTTCTGTGATGT
GGAAACCCTAGGCTCTGAGACACACTCTCTGGTGTCTGAGACAGAACCAAAGCAAT
AACGTTGTGATGCCCACAGGCCTGGAGCCAGCTAGCGACCTTGTGCCGCCCAGCTG
TCCATGGCCCGTGCAGAGCAGAGGACAGTGAGTGTCTGCACTGAGAACCTTAAACC
ACAGTTGAACATACCCACACCTGTTTGTCTTAAGCTATAGTGTA AAAACAAAGTTT
GGGCTCTGAAAATTTAACTGAAAAAGATTTCTTGT

~~Sequence ID 590~~ SEQ ID NO: 196

GTGGCAGCAGGCGCAGCCCAGCCTCGAAATGCAGAACGACGCCGGCGAGTTCGTGG
ACCTGTACGTGCCGCGGAAATGCTCCGCTAGCAATCGCATCATCGGTGCCAAGGAC
CACGCATCCATCCAGATGAACGTGGCCGAGGTTGACAAGGTCACAGGCAGGTTTAA
TGGCCAGTTTAAACTTATGCTATCTGCGGGGCCATTCTGTAGGATGGGTGAGTCAG
ATGATTCCATTCTCCGATTGGCCAAGGCCGATGGCATCGTCTCAAAGAACTTTTGA
CTGGAGAGAATCACAGATGTGGAATATTTGTCATAAATAAATAATGAAAACCTAAA

~~Sequence ID 591~~ SEQ ID NO: 197

CAGCAGCAGAAATGTTTGCAAGATAGGCCAAAATGAGTACAAAAGGTCTGTCTTCC
ATCAGACCCAGTGATGCTGCGACTCACACGCTTCAATTCAAGACCTGACCGCTAGT
AGGGAGGTTTATTCANATCGCTGGCAGCCTCGGCTGAGCAGATGCACAGAGGGGAT
CACTGTGCAGTGGGACCACCCTCACTGGCCTTCTGCAGCAGGGTTCTGGGATGTTT
TCAGTGGTCAAAATACTCTGTTTAGAGCAAGGGCTCAGAAAACAGAAATACTGTCA
TGGAGGTGCTGAACACAGGGAAGGTCTGGTACATATTGGAAATTATGAGCAGAACA
AATACTCAACTAAATGCACAAAGTATAAAGTGTAGCCATGT

~~Sequence ID 592~~ SEQ ID NO: 198

TACTCAATGAAAAACCATGATAATTCTTTGTATATAAAATAAACATTTGAAAAAAA
AAAAAAA

~~Sequence ID 593~~ SEQ ID NO: 199

nt:

565

CAGGATCAAGGTGAAAAGGAGAACCCCATGCGGGAACCTTCGCATCCGCAAACCTCTG
TCTCAACATCTGTGTTGGGGAGAGTGGAGACAGACTGACGCGAGCAGCCAAGGTGT
TGGAGCAGCTCACAGGGCAGACCCCTGTGTTTTCCAAAGCTAGATACACTGTCAGA
TCCTTTGGCATCCGGAGAAATGAAAAGATTGCTGTCCACTGCACAGTTCGAGGGGC
CAAGGCAGAAGAAATCTTGGAGAAGGGTCTAAAGGTGCGGGAGTATGAGTTAAGAA
AAAACAACCTTCTCAGATACTGGAACTTTGGTTTTGGGATCCAGGAACACATCGAT

Marked-Up Copy

CTGGGTATCAAATATGACCCAAGCATTGGTATCTACGGCCTGGACTTCTATGTGGT
GCTGGGTAGGCCAGGTTTCAGCATCGCAGACAAGAAGCGCAGGACAGGCTGCATTG
GGGCCAAACACAGAATCAGCAAAGAGGAGGCCATGCGCTGGTTCCAGCAGAAGTAT
GATGGGATCATCCTTCCTGGCAAATAAATTCCCGTTTCTATCCAAAAGAGCAATAA
AAAGT

~~Sequence ID 594~~SEQ ID NO: 200

CAGAAGAGTAAGCAAATCTCAAAGCAGCGAAAGGGAAGAACTAAAAAAGGTAGAG
CAGAAATAAGAGAAAATAGAGAAGAGAACAATTGAGAAAAATAATTGAAACCAAAA
GGTGGTTCTTTGAAAAGCCTAACAAAATGGACACATCTTTAGTTAGAGTGACCAAG
AAAAAAGGGCAGTGACTCAGATTACTTCATTCAAGAGTGAAAGAGGGGCACATCACT
ACCAATTTACAGAAATAAAAAGGATTATGAGGAAATACTACAGATAATTGATGACA
TTAACTTAGAAGAATATATTTCAAGAAAGACACAACTACTGAAACCGACTCAAGA
AGAAACAGAAAAATCTGAACAGACCTATAAAAAATAGAGATTTAATTGATATTCAGA
AAGTTTCCCAAAAAGAAAAGCACTGGCCAAGATGACTTCACTGGTGAATTCTATCA
AGTGTCAAAGATGAATTACTGACATTCATTACACTCCTTTAAGAAATAGAAGAGG
GGACATCACTTTTCAAAGCATCGACATTCTAATCATTAGTCCCTTGGTTTCCTGCT
CCCAAAGCCAGGTGATGTATCACAAAAAACCCTACAGACCCACTGGGCACAATG
GCTTTATGCCTAT

~~Sequence ID 595~~SEQ ID NO: 201

nt:

98

CTTTGCTCGAATNGTCAGATAAGGATTCTGTGAANGGAGATGAGATTTCCATCCAT
GCTGACTTTGANAATACATGTTCCCGAATTGGGGNCCCCAAA

~~Sequence ID 596~~SEQ ID NO: 202

CTCAAGTGTTCCCTCAGCTTAGGCTTTGTTTAAATGATCCCACCCAGGGGCGATGG
TAGGGAACAACAGGGTCACTAACTATTTGGCTGGCTACAACCTCTGGGAAATGGTA
AGACAGGGAAAGGCCATGTTGTTTCATTCCCTTGTGCAGATCTAGGGAGAACCGCAG
AGAGAACAGTTAGCATTTCTTGTTCAATGAATTATCCTATTAAGAACACTGGATGT

~~Sequence ID 597~~SEQ ID NO: 203

CGGNCGCGGTGACGCTACTCCTACCTATCTCCCTTTTATACTAATAATCTTATA
AAAAAAAAAAAAAAAAAAAAAAAAA

~~Sequence ID 598~~SEQ ID NO: 204

nt:

362

GGCATGTGCCTGTAGTCCTAGTTGCTGAGGTAAGAGGATTGCTTGAGCCCAAGAGT
TCAAGGCTGCAACAAGCTTTGATTGCGCCACTGCACTCCANCCTTGCGGACAGACT
AAAACGCTGTCTCAAAAAAAAAACAAAAACGACNAAAAAAAAACAAAACAGAAAAA
ATTAACCTAGGCAATGACAGTCCCTGGCAAATGCTGGGAGGGAGGCAACANTGGTC
AAGGAAGGTAACCCTGAANCAGGACTTGTAAGCAAATAANATTGGGAGGCCAAGG
TGGGTGGATCACNAGGTCAGGAGTTCGAGACCAACCTGGCCAACATAGTGAAACCC
CGTCTTTCTAAAAATACAAAAAATT

~~Sequence ID 599~~ SEQ ID NO: 205

GACAAAAGAACCATTTGGATACATAGGTATGGTCTGAGCTATGATATCAATTGGCT
TCCTAGGGTTTATCGTGTGAGCACACCATATATTTACAGTAGGAATAGACGTAGAC
ACACGAGCATATTTACCTCCGCTACCATAATCATCGCTATCCCCACCGGCGTCAA
AGTATTTAGCTGACTCGCCACACTCCACGGAAGCAATATGAAATGATCTGCTGCAG
TGCTCTGAGCCCTAGGATTCATCTTTCTTTTCACCGTAGGTGGCCTGACTGGCATT
GTATTAGCAAATCATCACTAGACATCGTACTACACGACACGTACTACGTTGTAGC
TCACTTCCACTATGTCCTATCAATAGGAGCTGTATTTGCCATCATAGGAGGCTTCA
TTCACTGATTTCCCTATTCTCAGGCTACACCCTAGACCAAACCTACGCCAAAATC
CATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCACAACACTTTCT
CGGCCTGTCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACACCACAT
GAAACATCCTATCATCTGGAG

~~Sequence ID 600~~ SEQ ID NO: 206

nt:

595

TTCAAATTCTTGNTAANAGTCTTTGTTCTGAATTTTACTTTGTCTGTTATTCCTAT
AGCCTTTCCAATTTTCTTTTCGCTTGGATTTTACGTGATAAGTTTTTCCCCCATTT
TACTTTTANCAACTCTATATTTTTTTAGTTGAGGTTGGGTTTCTTGTAACAGCATA
TAATTTGGGTTTTTTAATCCAATCTGAAAATTAATGTCCTTAATTTTGTGTTTATA
CCATTTACACATAATGTACTCATATATAAGGTTTAACTGAAACCTACTATCTTGCT
AGTTGTGCTCTACTTGAATTTTTTTTTTAGTATTCTGTTTTAATTGACCAACATTTG
ACTGTATCTCTTTGTGTAATTCTTTTACAGGTTGCTGTAGGCATGACAATATATAC
ACTTAACTTTTCTCAGTACACTGAGAGTTGAAATTGTAGTACTTCGAGGAAAACAT
AGAAAACCTGCAATGATATCGGTTACATTTTACCACCTCCATATGTTGCAATTATT
AAATGTATTAGATCTGCCTACCTCGAAAACCCATCAGTCTTTTAACTTTGCTCTCA
ATGGTGATTCATATTTTTTAAAAAACTTGAGGCAA

~~Sequence ID - 601~~ SEQ ID NO: 207

nt:

522

TCGACCGGGTTTGGAGCAGTGCCTTGTTTGCTGTGCAGCGGATACTCTACAGGTAC
ATTTTCCTTTTTTGAACCAAAAGGGAGGGATTTGACAATATTGATGGTAGATCTTTT
TTCTTTAGCAAGAATTAAGGATTTTGGTGGGTGGGGGGAGGCTTCTGTGGGGACCA
AGACAATGTACTGTCAGTCAGGATTTAAGTCGAACCTACCTCATCCCTTGCCCCAGA
GAACAGTTGATCGTGTTTTAAACCAAAAGGTGCGGAATGGAGAGAGGGAGGCGGTG
CATTGCAGCTTCCGATAGAGCTTTTTATTTTTGGATATCAGGAACCAATTTTGAAG
ATTTCTTAAGAAAGTCATTTACATCAGGGACATGAAGAGCAAAGTAGGTATTTTTG
GTCAGTACTTGAATTTGATAGGCTTTATGCAAACAACTCTCCCTCTGCTGGAGTCT
GGCAAGTTTGCTTTTCACTGGACGCTAATTCAAGTGCCATACAAAATAAATAAN
AGTTTTACTTATAACACA

~~Sequence ID - 602~~ SEQ ID NO: 208

CAGAAATCGCAATTGAAGACCAGATTTGTCAAGGTTTGAACTGACATTTGATACT
ACCTTCTCACCAACACAGGAAAGAAAAGTGGTAAAATCAAGTCTTCTTACAAGAG
GGAGTGTATAAACCTTGTTGTGATGTTGACTTTGATTTTGCTGGACCTGCAATCC
ATGGTTTCAGCTGTCTTTGGTTATGAGGGCTGGCTTGCTGGCTACCAGATGACCTTT
GACAGTGCCAAATCAAAGCTGACAAGGAATAACTTTGCAGTGGGCTACAGGACTGG
GGACTTCCAGCTACACACTAATGTCAATGATGGGACAGAATTTGGAGGATCAATTT
ATCAGAAAGTTTGTGAAGATCTTGACACTTCAGTAAACCTTGCTTGGACATCAGGT
ACCAACTGCACTCGTTTTTGGCATTGCAGCTAAATATCAGTTGGATCCCACTGCTTC
CATTTCTGCAAAAGTCAACAACCTCTAGCTTAATTGGAGTAGGCTATACTCAGACTC
TGAGGCCTGGTGTGAAGCTTACACTCTCTGCTCTGGTAGATGGGAAGAGCATTAAAT
GCTGGAGGCCACAAGGTTGGGCTCG

~~Sequence ID - 603~~ SEQ ID NO: 209

nt:

624

GACACACGAGCATATTTACCTCCGCTACCATAATCATCGCTATCCCCACCGGCGT
CAAAGTATTTAGCTGACTCGCCACACTCCACGGAAGCAATATGAAATGATCTGCTG
CAGTGCTCTGAGCCCTAGGATTCATCTTTCTTTTACCAGTAGGTGGCCTGACTGGC
ATTGTATTAGCAAACCTCATCACTAGACATCGTACTACACGACACGTACTACGTTGT
AGCCCACTTCCACTATGTCCTATCAATAGGAGCTGTATTTGCCATCATAGGAGGCT
TCATTCACTGATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCTACGCCAAA
ATCCATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCACAACACTT
TCTCGGCCTATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACACCA

Marked-Up Copy

CATGAAACATCCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAGTAATATTA
ATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAAGCGAAAAGTCCTAATAGTAGA
AGAACCCTCCATAAACCTGGAGTGACTATATGGATGCCCCCACCCTACCACACAT
TCGAAGAA

~~Sequence ID - 605~~ SEQ ID NO: 210

nt:

338

ACCTGAGGCCTCGGTGGGGCCAGTGCGACGCTGGCTTAAGGAGCTGGAGGGGTTC
TAATACACATTTAATTCAGTTTCTCTTCCCTAAGAGGCTGCCGGAGTTGGGGCCTC
CTCCAGCAGAGACCCTCGGACCCCTGCAGGGCCTGGACTTGGGGTGAACAGGGCTT
CAGTCAGCGCAAGTATTCATTTGCATTTGGTAATTTTTCATGCCACCTATTTATG
AATATATAAATCTTTATACCAAATCTATTTTTTAAACATGGAAAAGTTGCCTTTA
TGGAACCTTGGCAGAGCCAGAGTGTACACATTCCTAAACCATTAAACAGATTTCTA
TA

~~Sequence ID - 606~~ SEQ ID NO: 211

nt:

556

GGATAATGATACCTCTGACCTTTCTTCCTTTTGGGAAGTACTTGAGTGTGCAGCTG
CATGAGGCCTCAGCAGGAGAGAGATTTTAGGTCCAAGAAGCTATACCAGTAGGACA
AGGCAGGAAAATACTACACTTTCAGGATCAAGCCCCTCTGACTCTCATTTGGAAAC
TGGATGTTTGCTAAGCACCTGCTTCTTAAGGATGCCGAGGGATTTAATGATACTCC
CAGAAACCTGGAGAGATTAATGGGGCCTATGGAGAAGTGCTCTGAACTCAGTGTTG
GGACTTGAATAAAATTAACCATTGTTCATGTTTTTCAGAACAACTAAGCTGTTTTATA
TTTCATGTGCATGAAAGCCCTAGAACTAAGTTGTGTTATTTCCAGAAATGAAATAG
ATCCCACAGTTAGATGATGTGGCCATTAGGAAGTACCAAATTTATAAAAATCACTG
GAGGTCTGTCTGAGCAGTACCTAATAAAATATAGTATACTGAAAGTGAACAGATCT
TTGTCTCTTTCTTTGGCTGCTTGATACTTTATCTGTGTCTGCCGGACAGTGC

~~Sequence ID - 607~~ SEQ ID NO: 212

CAATAAAAGCAGGTTAACCTCAATGATAGCAGTTAAATGTTCTATCTTATGTATT
TCTTTTAAAGTATTACCATTATGGTGCTACTGAGCGTTTTCTTTTGGTAAAAAGAAA
AATGCCATGGGCTGCAGTCTTCTTCCATCACTTTTCCCTACCAGGTCCATTAATAT
GCTTATAACACTAGTGCCAGTTATTTTATTTGATAATGCTTATGGTATTTGTATAT
TTGTTTGCATTCCAATTTTGTTTAATAATGAGTGTGTAACTGCATACGTAAATA
AATGTAAATACTAATGTACTGCTGC

~~Sequence ID 609~~ SEQ ID NO: 213

TTTTATTACCCAAGTTTTTAACCTCTGTCTGGTGATTTGTTGTTGTTGTTGTTGTNG
TTGTTGTTGAAGTTCAGGCTGCATGTGGGATAGGTTTGCTCAGGCATACTTCTTAG
GAAGTAGTCACTTGCATGACTGTTTTTTGGGATAACTCTTTGAGTATTTGGAGAGGT
CTATTGTAACCTTCTGAAAGGCATTGTTTTTACGTATGAATGTTCTAAAATTCATTC
TAAATGGTCATGAAAAGAAAAGGATTCACATTTTAGAATGGCAATAGTCCCTGAGG
ACTATTATGTCTTTTAGATTTCTGTGGGTTTCTAGGAATGTTAGTGTAACCTTANA
TTTCCACCTACCTGATTTCTGGATGTGCCTATTGGAACCTGCTGAGATCTTTTTTTT
TTCCTTAACATGTTGTCCCCCTTGACCCGTA CTTCGAAACTAAACATATTATTTTAT
TTGCTTACACTTCAGGAGGCAATTGGCAGACACCAGGCCAACAGTCT

~~Sequence ID 610~~ SEQ ID NO: 214

GCTCTGACCCCAAGTTGGAATGTATCTGTACTTTGTCCGGCTTCCACTCAAGGACC
ATTTATGACATTGCTTGGTGTCAGCTGACAGGGGCTCTGGCCACAGCTTGTGGGGA
TGACGCGATCCGCGTGTTTCAGGAGGATCCCAACTCGGATCCACAGCAGCCCACCT
TCTCCCTGACAGCCCACCTTGCATCAGGCCCATTCCTCAGGATGTCAACTGTGTGGCC
TGGAACCCCAAGGAGCCAGGGCTACTGGCCTCCTGCAGTGATGATGGGGAGGTGGC
CTTCTGGAAGTATCAGCGGCCTGAAGGCCTCTGAGCTACCTCGACTTTGGACAGAG
TAATGACTCCCCAGAAAACGTCATATAAGACTTTACCAGCCCCTGAGAGGACCAGG
AGGAGCATCCTTGACCTTCATTTAACTTGGCTCACTTCTCTTCANACTTGGGTAGA
AGTGCAGAGCCACAAAATTGCTTTCCTTCCCCGCCTTTGACATGAGGCCTTCAGTA
AAG

~~Sequence ID 611~~ SEQ ID NO: 215

TGCAGGATCCGTCGACT

~~Sequence ID 612~~ SEQ ID NO: 216

nt:

576

GAGAAATATAAGATTATGTATAGATCAAATCTACCTCTATTTGGTGTCCTGAAAGA
GATGAGGAGAATGGGACAACTTGGAAAGCTTATTTCAAGATAACATTCCTGAGAA
CTTCCCCAATCTTGCTAGAGAGGCCAACATTAATAATTCAGTAAATGCTGAAAACCTC
CAGTAAGATATTTCTTAAGAAAATTATTCCCAAGATATATACTCATCAAATTATCT
AAGGTCAAATGAAGGAAAAAATTTTATAGGCAGCTAGAGAGAAATGTCAGGTCACC
TACAAAGAGAATGGCATAAGACAAAAAGTAGAACTCCCAGCAGAACTCTAAAAGC
CAGAAGAGATTAGGGGCCAATATTTAACATTCTGAAAGAAATTCCAACAAGGAATT

Marked-Up Copy

TCATATCCAGCCAACTAAGCTTCATAATTGAAGGAGAAATAAGATATTTTCCAGA
CAAGCAAATGCTGATGAAATCCATCACCACCAGACCTGCCTTATAAGAGCTCCTGA
GGGAAGCACTAAATATTGAAAGGGAAGAACTTTATGAACCATTTCAAAAACACATT
TAAGTNCACAAAGCAG

~~Sequence ID - 613~~ SEQ ID NO: 217

nt:

341

CCTTATTTTACAGGTGAAAAACCACGAATCAGATAGATTTTATTTGCCCAAGTCA
CATAATATTAAGAACAGGCCAAGTGTGGTGGCTCATGTCTGTAATCTGAGCACTTT
GGGAGGCTAAGGCGGGTGGATTTCTGAGCCTAGGAGTTTGAGATCAGCCTGGGCA
ACATGGCGAAACCTCATCTCTACAAAACATACAAAAATTAGTCAGTGTGGTGGTGA
GAGCCTGTAGTCCTGGCTACTCGTGAGGCTGAGGTGGGAGCATCACCTGAGCCTGG
GAAGTCGAGGCTGCAGTGGCAACAGAATGGGTAACCTGGACATCAGAGTGAGACCC
TGTCT

~~Sequence ID - 614~~ SEQ ID NO: 218

CTCACACCTGTAATTCCATTACTTTGGAAGGCTGAGAGAGGAGGATCAGTGGAGCC
CAGGAGTTTGAGACCAGCCTGGGCAATATAGGGAGACCCTGTCTCTACAAAAATGA
AATAGCCAGGCGAGGTGGCATGTGCCTGTGGTCCCAGCTACTTGGGAGACTGAGGT
GGAAGGCTGCCTTGAGCCCAGGAGTTCCAGGCTGCAGTGAGCCATCATTATGCCAC
TGCACTCCAACCTGGGAGACAGAGTGAGAGAGACCCTGTCTCAAACAAACAAACCC
AAAATAGGCCAGGCACAGTGACTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGA
AATAGGCGGATCATTTGAGGTCAGGAGTTCAAATTCAAGACCAGCCCGGCCAACAT
GGCAAAACCACATCTCTACTACAAATAAAAAATTAGTTGGGTGTGGNGGAGCATTC
CTGTAATCACAGCTATTCAGGAGGCTGAGGCATGANAACCGCTTCA

~~Sequence ID - 615~~ SEQ ID NO: 219

nt:

379

TAAATTTAAACATTTTAATTAGCTGGCATGATGGCATGCACCTGTAGTCCTACCT
ACTTGGGAGGCCAAGGCAGGAAGATTGCTTGAGCCCAGGAGTTTGAGCTTACTGTG
AGCTGTGATCACACCACTGCACTCCAGCCTGGGTGACAAAGGAAGACCGTATTTCT
AAAAAATAAAAAATACAAATACAACTACAACTAGCACTAGACCAACAGTGACTAT
GTACCATGAACTGAGGAATATTATTAATTCACCATTTGCATCTGAGGTTAACAAT
ATGTCAATGACTTAAATAACATCATATCTCTGAGAGTAATTTCTCCTATATTTCCA
TGACAAATGTTAGATAATTTTCCATTTTTTCCATTCAACAAAA

~~Sequence ID 617~~SEQ ID NO: 220

TTTTCAGGCATGTCAGAGAAGGGAGGACTCACTAGAATTAGCAAACAAAACCACCC
TGACATCCTCCTTCAGGAACACGGGGAGCAGAGGCCAAAGCACTAAGGGGAGGGCG
CATACCCGAGACGATTGTATGAAGAAAATATGGAGGAACTGTTACATGTTTCGGTAC
TAAGTCATTTTCAGGGGATTGAAAGACTATTGCTGGATTTTCATGATGCTGACTGGC
GTTAGCTGATTAAACCCATGTAAATAGGCACCTTAAATAGAAGCAGGAAAGGGAGACA
AAGACTGGCTTCTGGACTTCCTCCCTGATCCCCACTCTTACTCATCACCTGCAGTG
GCCAGAATTAGGGACTCAGAATCAAACCAGTGTAAGGCAGTGCTGGCTGCCATTGC
CTGGTCACATTGAAATTGGTGGCTTCATT

~~Sequence ID 618~~SEQ ID NO: 221

nt:

598

GATTAACTTTCATTTTAAGCTCTTCTCTACTAATTCTGTTTCGTATGTTTATTCATT
TTGCGTTGATCATATTTTGTACACCAGGCACTCTTCTCAGTTTTATATGTGTGTTA
ATTTACTCCTTTCAAGAGCCCTATGATACATGAATTTATCTCCATTTTATAGATGA
GGAAATTAAGACCTAGAGTTACTGAACTTGCCCAAGGTTATACAGCTGATGGGTAG
GGCCAGAACTTTGCCTCAGAGAATCTGAATTTCCAAAAAATAACCTAAAAGAGAAA
TTTAAGTACTAATTAGTAAGCAAAGAAATGCACATTTAAGGAAGACAGTGCACATT
TAAGGAAGACAGTAACCTTTTATCTATTAGAGAAAAACACACATTCTGTCTTTAAC
ACACACATAAATCTTATATTGGCAGGGATTTTCTTTATTCAGCAATTATTTATTGG
TTGTCTGCTTTGTGGTACACATAAATGCTGGGGATAAACACTTAATAAAATATACT
TCCTTCTCTTGAATATCTTGCACTTTAAGTGGGAAGGTAAGTCAACAGAGTAGAGG
TGATATATCCAAGTGATAGACTGTTTCATTGCCAGTAG

~~Sequence ID 619~~SEQ ID NO: 222

GTTGCCTGAGAGTGACCTTTGCATCTGCCTGTCCAGCCAGCATGGAACCAAAGCGG
ATCAGAGAGGGCTACCTTGTGAAGAAGGGGAGCGTGTTCAATACGTGGAAACCCAT
GTGGGTTGTATTGTTAGAAGATGGAATTGAATTCTATAAGAAGAAAAGTGACAACA
GCCCCAAAGGAATGATCCCGCTGAAAGGGAGCACTCTGACTAGCCCTTGTCAAGAC
TTTGGCAAAGGATGTTTGTGTTTAAGATCACTATGACCAAACAGCAGGACCACTT
CTTCCAGGCAGCCTTCCTGGAGGAGAGAGATGCCTGGGTTTCGGGATATCAATAAGG
CCATTAAATGCATTGAAGGAGGCCAGAAATTTGCCAGGAAATCTACCAGGAGGTCC
ATTCGACTGCCAGAAACCATTGACTTAGGTGCCTTATATTTGTCCATGAAAGACAC
TGAAAAAGGAATAAAAGAACTGAAT

~~Sequence ID 621~~SEQ ID NO: 223

Marked-Up Copy

TGGTACTGAACCTACGAGTACACCGACTACGGCGGACTAATCTTCAACTCCTACAT
ACTTCCCCCATTATTTCCTAGAACCAGGCGACCTGCGACTCCTTGACGTTGACAATC
GAGTAGTACTCCCGATTGAAGCCCCCATTCGTATAATAATTACATCACAAGACGTC
TTGCACTCATGAGCTGTCCCCACATTAGGCTTAAAAACAGATGCAATTCCCGGACG
TCTAAACCAAACCACTTTCACCGCTACACGACCGGGGTATACTACGGTCAATGCT
CTGAAATCTGTGGAGCAAACCACAGTTTCATGCCCATCGTCCTAGAATTAATTCCC
CTAAAAATCTTTGAAATAGGGCCCGTATTTACCCTATAGCACCCCCTCTACCCCCT

~~Sequence ID 622~~SEQ ID NO: 224

TTTTTCTTGTTTTTGTGTGTCTACCTTGGCATATACTAAAGGAAGGTGTGTATTCA
TTTATTACATGATATCTCTGGGTTATAATTATTTACATATATGAATTTGAAAGAAA
GATTGAGAGGGATATGTGTGACCTTTGTTTCATTATGATCATTTACATGACTAAAG
ATAAAGATCATATGTCTGATTTTCAGTTTAATGGCAAGTTACTTAAAATAAATGAA
ATATGTTTTTATTGTTTTTCGTGGGTTTGATGCTTTGTGTTTTATTTCAAGTAACCT
GAGAATGCATTGTGTTTGGTACTGTTTTTTATGAATATCATTAATAAATTTATTTAA
GGAGAGAGTAATTTTGCAATAATATTTTTGATTTATTTGAAAATAAATTTCAAGAT
AAATGAAATAATTGAAATTTTCTAAAGAAGGAATTGAATATATTTTTACATTTGAA
TGAACTAAGGATTAACCTGAACCATTTATATATAGTACTTTCAGAACTGAATGTCTT
AAATGATAAAGCTCTAATTGGTTAAAGTGACTTTCTTTCAAGTCAAAGAACCCAGA
AACTGAATAGATGATCTAACTACTGCCACTGAGGTTTTGGATTAGTGAGTATAAAT
TT

~~Sequence ID 624~~SEQ ID NO: 225

TGCAGGATCCGTGCGACT

~~Sequence ID 625~~SEQ ID NO: 226

GACAATCAGAGCAGATCTTGGGCTTCTGTGGCTCATCTCAGCCCTTTATAACTGGC
CTGAGAAGAGGGTTTATCTACTTGTGCAAGTGGCCCAGAAATCTCACTCGTACATG
AGGCTTTGGAACATCCTTGCAAAGGTACGCTGAAAGCAAATTGCTGTTTTCTGGT
GGTTCTGCACGTTTCCTAACTTTTATCATAGTTTGATTTTCATTATTTAAGAAAAA
ATAAAAAATCCAAAGACCATAAGATGGCATTAGATTTTTTTACCATTAAATTATTAA
TGCCTATTTGGTGCTCATAAAGATTAATCATGTCACGCATGTTTCCAATCTTTCTT
TTGCAGTATATTATTTTCTAAAAATGTTTACATGCAAATTTAAACCAAGATTTATC
AGTA

~~Sequence ID 626~~SEQ ID NO: 227

TTGGAAGAAATAAACCAAGGCAGAAAAATTTTAAATGGCCAAAATAAATTGTATTG
CTAACTTAGATGGCCACAGATGGGGGCAGGGGTGGAGAGAGGAGAAATTGAAAACN
CCACAAAGACCCCGCAATGGCTAGAACTTGAAATCTCTGGATATTGCAACAATAGC
AGCCTCCTTAAGTCAGCAAAAAGATAAAGATTGATCCAATGTTCTATATTACAGAA
CAGAGCAGATTGTCAATATAGCAAATAAAGTTACCGTTGAGTGGACTGCGCTGTNT
AAGCTGCTTGGTTGGCCTTAAGTGCCGACAATTAAGAGATGAAGGCAATGAGAACT
GAAACAAACATTTAAGTTCAAGACCCAGTTTACTGACACTGGGACTATTACTATAT
CTCTTTGGGCCTCAGTTTACTTATCTGTAACATTAAGAGGTTGGATTACATGATGT
CTCACGATTCTTTTTTTTTTATTTAGAGATGGGGTTTTTGCTCTGTTGCCCAGGCTGG
AGTGCAGTGGCATGATCATAGCTCACAGCAG

~~Sequence ID 627~~SEQ ID NO: 228

CCAGCCTGTCACTGGCCTGGCCAAGGAGGAGAGACAGGCCAGGGATTCTGGTCCTA
ACTCTACTGGCCACACTGTGTGGCCTGAGACCCCCCTTTCCCTCCCAAGCCCCTGC
CTCCGCATCTGCGTGGTGAAGGCCATTGGCCCTCATCGGTGGATCTGCGTTTCCTC
GGGCCTACACTGTCTAGGATTGTGCGGGGCTGGTGAGAGAACAAGATCTCTTCCGT
GTTCAAGGCAGACTTCCTGCCCCCTGCACCCTGCTCTCTCCAGGCCTTGAGGTCA
GTGTGAGCCCCAAGGGCAAGAACACTTCTGGAAGGGAGAGTGGATTTGGCTGGGCC
ATCTGGATGGAAGGTAAAAAAGAAAATCCCTTGAAAGGAGATTGAGGGAAGTTT

~~Sequence ID 628~~SEQ ID NO: 229

nt:

419

AAGAGAAAGGACTCAGTGTGTGATCCGGTTTTCTTTTTGCTCGCCCCCTGTTTTTTGT
AGAATCTCTTCATGCTTGACATACCTACCAGTATTATTCCCGACGACACATATACA
TATGAGAATATACCTTATTTATTTTTGTGTAGGTGTCTGCCTTCACAAATGTCATT
GTCTACTCCTAGAAGAACCAAATACCTCAATTTTTGTTTTTGAGTACTGTACTATC
CTGTAAATATATCTTAAGCAGGTTTGTTTTCAGCACTGATGGAAAATACCAGTGTT
GGGTTTTTTTTTAGTTGCCAACAGTTGTATGTTTGCTGATTATTTATGACCTGAAA
TAATATATTTCTTCTTCTAAGAAGACATTTTGTTACATAAGGATGACTTTTTTTATA
CAATGGGAATAAATTATGGCATTTTTT

~~Sequence ID 629~~SEQ ID NO: 230

CTGAGAGTCACTGTGTTTTTAGCCAAATCTAAGGGAGAAAATGAATATTGATAGCA
GCATGCTGTAGCCAGCTCCTTAAAGGAAGGATGGTGCCTGGTACAGAGTTAGAGTT
AGTGCTTCAGTAAATAATGAATGTGTGCTAGGTAGGTTCTGCTGGGTAGGCTGCAT

Marked-Up Copy

GCATTGACCAATTTATTCTCCTTGTTTCAAACAGGATTTAAGGGCACTTATATA
TATATATTTTTTAGTTTTTTAATGTAAATGAGAGAATAAAGATATATATATATGT
CTATATATGTATATATGTATATATATGTCTATATGTCTATATGTATATATGTCTAT
ATGTATATATGTGTGTGTGTATATATATATATATATATATAAGTTTTCTGTTGCTA
GCATAACAACTACCAGAACTTAGCAACTGAAACAACATGAATTTATCTTACGGT
TCTATAGTTCAGAAGTCTAACGTGTCACTGGGATGAAATCCAGGTTTCAACAGGAC
TGGGTTCCCTTCTAGCTCATTACGTACCTGGCTCATTACAGGTTGTNGGCAGAATA
TACTTCCATGAACTGTAGGGCTGAGACCCCGTTCCTTCCTGGCTATCATCTGAAA
ACTTTC

~~Sequence ID 630~~ SEQ ID NO: 231

AGGCGCAGCCCAGCCTCGAAATGCAGAACGACGCCGGCGAGTTCGTGGACCTGTAC
GTGCCGCGGAAATGCTCCGCTAGCAATCGCATCATCGGTGCCAAGGACCACGCATC
CATCCAGATGAACGTGGCCGAGGTTGACAAGGTCACAGGCAGGTTTAATGGCCAGT
TTAAACTTATGCTATCTGCGGGGCCATTCGTAGGATGGGTGAGTCAGATGATTCC
ATTCTCCGATTGGCCAAGGCCGATGGCATCGTCTCAAAGAACTTTTGACTGGAGAG
AATCACAGATGTGGAATATTTGTCATAAATAAATAATGAAAACCTAAAAAAAAAAAA
AAAAAAAAAAAAAA

~~Sequence ID 631~~ SEQ ID NO: 232

TNCACTCACACACTCCCAAACCTTAACAAACACATACATGTGCAGCCAACCCAATG
GGCCAGCCTCTTTTATGCTCCTCACATGTTTCCTTTAACTGGAATACCCATGACAG
CTCCCTACATAGTTACTTGTAAACTCCTCCTCTCTGTATAAGTTTTCTGAATTTT
TTTGATAAAATTAAGTTGTGCCACCCCTTTATGCTCTCTTANAACCTTTGTTCTGTT
CTCATGGCTGTTCTGCAACGAATCTCATTGTGTTCTCCTACTCAATTACATTCCTG
CGTCTCCCACTAGATGGCAGACTCTTTGAGAGTAGGAGATTCCCTTGTTATCTCTG
GATCCCTGGCACTTGCAGAAAGCCTGTTACGTAATAATTGCTCAACAATTAGTTTT
TAAATAAATGAATTATTTTTTAAACGCCAAAATTACAATGATTGTGCATTAAGTGA
AAGATGACCATCTAAAACATAAAGCCATGCTTCATGACATTGGC

~~Sequence ID 632~~ SEQ ID NO: 233

GACCATTCAGGGAAATTTTATAAAAAATGCAGATACTGTCTTGAGCAGATCGAAAT
GCCGATGAGGTGGATGCAATTTCTTTTGTGCAAGCAGTGCACGGTGCCCCCCCCCT
CGGGTGTCCGTGCTGTGCCTTAGCTTCCCCAGGTGCCGGGACTCACACCTGCTAGG
GGCTGGGCAAGGCCCCGGCTCTGCTTTCTCTGAAGGGCTTGTCCAAGTTTCATTGCC
CTGTTACAGGTGGTCAAGACGTCCGGCCGCCTTGACCCAGGCTACCCTTAGCCAAT

Marked-Up Copy

ATCCTCTGCCCCCTGGGTGGTTGGTGGCTGGGCCTCAGGGTGGGCAACGTTAGGGGT
TTGGCGAAAGCCCCGCCCATGGGATTGAGGGACGGGGCTGCACTCCAACCGTCTGC
ACCTGCTCTTCCCCCACCCCTGTGGGACCTCATCTTCACGTGCCATGTGTGCTGAA
GGCCCAGGGCCCAGCAGGGGGCAGTGGCACCTGTTGACGGAAAAGCCGAGGTGCTT
ACCAATGGACCTTCTGGCCCCGCCCTCCCCTGTACTTGTCTGGGCATTGAGGGCCCCG
ACCTGTGCCTACCCGCA

~~Sequence ID 633~~ SEQ ID NO: 234

CAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCTGACTCCTGAGG
AGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGT
GAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTC
CTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTC
ATGGCAAGAAAGTGCTCGGTGCCCTTAGTGATGGCCTGGCTCACCTGGACAACCTC
AAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCC
TGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTG
GCAAAGAATTCACCCACCAAGTGCAGGCTGCCTATCANAAAGTGGTGGCTGGTGTG
GGCTAATGCCTGGCCCCACAAGTATCACTAAGCTCGCTTTCTTGCTGTCCAATTTT
TATTAAAGGTTTCCTTTGTTCCCTAAGTCCAACCTACTAAACTGGGGGATATTATGAA
GGGCCTTG

~~Sequence ID 634~~ SEQ ID NO: 235

nt:

511

TTTTTTAATTTACCAAAATTTGTTGACGTCCCTTGATTTGCTGATAGGGACAATA
ATTAAATATTTTCCACTTGTTTTTATAAAAAGTGAATGGTGATTTGTTTAACAGA
TGTTGACTTAGCACCTTCTCTTTTTTTTTTTTTTTTTTTGAGTTGGAGTCTTGC
TCTGTCACCCAGCTGGAGTGCAGTGGCACGATTTTCGGCTCACTGCAACCTCCGCCT
CCCAGGTTTCGGGCGCTTCTCCTGCCTCAGCCTCCCANATAGTTGGGATTACAGGTG
CATGCCGCCACNCCTAGCTAATGTTTTTTGTATCTTGGTANANATGGNGTTTCACC
TTGTTGCCCATGCCGCTCTTGAACCTCCTTGGCCTCCCAAAGTGTTAGGATTACAGG
CGTGAGCCACTGTGCCTGGCCCCAATTTANCACCTTACTGGGTGCTGAGGCTGTGA
GCCATAGTAGAATGCATGTGATCCAGGGCCTTGCTGAATTCATGGGCTAATAGGGA
GCCTGAC

~~Sequence ID 635~~ SEQ ID NO: 236

nt:

592

TGAGCGTTGGGCTGTAGGTCGCTGTGCTGTGTGATCCCCAGAGCCATGCCCGAGA

Marked-Up Copy

TAGTGGATACCTGTTTCGTTGGCCTCTCCGGCTTCCGTCTGCCGGACCAAGCACCTG
CACCTGCGCTGCAGCGTCGACTTTACTCGCCGGACGCTGACCGGGACTGCTGCTCT
CACGGTCCAGTCTCAGGAGGACAATCTGCGCAGCCTGGTTTTGGATACAAAGGACC
TTACAATAGAAAAAGTAGTGATCAATGGACAAGAAGTCAAATATGCTCTTGGAGAA
AGACAAAGTTACAAGGGATCGCCAATGGAAATCTCTCTTCCTATCGCTTTGAGCAA
AAATCAAGAAATTGTTATAGAAATTTCTTTTGAGACCTCTCCAAAATCTTCTGCTC
TCCAGTGGCTCACTCCTGAACAGACTTCTGGGAAGGAACACCCATATCTCTTTAGT
CAGTGCCAGGCCATCCACTGCAGAGCAATCCTTCCTTGTCAGGACACTCCTTCTGN
GAAATTAACCTATACTGCAGAGGTGTCTGTCCCTAAAGAACTGGTGGCACTTATGA
GTGCTATTCGTGATGGAGAAACACCTGACCCA

~~Sequence ID -- 636~~ SEQ ID NO: 237

nt:

572

CTTANAAGAGTTGCTCATTACACCCACGCCCTTGCCCAAGGCTGGCCCACTCAGA
GCGAAACTTAACTTTTGTCTGGATGGGAAGAGAAGTAAGTCTACCCCGAGGTTGCC
ATGTTGAAGAGTGAGAGGTCCAAGTGATTCTGTGCATTGAAACCAAGACACCCAC
CCAGAACACTTCTTCCCTCCCTCAGCCCAAACCAAAGGCTGGGGTTCTCATCTCCA
AGTGGCTGTTCTCCAACTTTCCCAAGCCGCTTGCAATCCCCAGACTGGACTACTGT
GGCGGTTAGGTTAGATTTGAAGACGGGGCCCAGGCTGGGTATGAACGGGTGCAGCC
CTCTTCTCCTCTTCCCCCCCACATCTCTCATGAGAGAGGTAGTGGCATTTCCTTCT
CAGGGAGCTTCAATGGGAAAGGTCTCGAAAGCTTCAGGAGGAGCAGAATACCAACG
CAGGGGGATGGCTGTAACGATCTCACCGTCTCCTAACCTCAGTCCCTTTTTTGAGA
GTGAATGGTGGAGGGTGGGAAAGGGACCCAAATTTGTAGATCTCTTTGTCTGGGGG
AGGGGAANGATG

~~Sequence ID -- 637~~ SEQ ID NO: 238

nt:

482

TTAAAACAGGCGCAGGGGTAAAAATGAGAATGAATCTGAAAAAGAGAGTTGGTGT
TTAAAGAGGATGGACAAGAGTATGCTCAGGTAATCAAAATGTTGGGAAATGGACGA
TTGGAAGCATTGTGTTTTGATGGTGTAAGAGGTTATGCCATATCAGAGGGAAATT
GAGAAAAAGGTTTGGATAAATACATCAGACATTATATTGGTTGGTCTACGGGACT
ATCAGGATAACAAAGCTGATGTAATTTTAAAGTACAATGCAGATGAAGCTAGAAGC
CTGAAGGCATATGGCGAGCTTCCAGAACATGCTAAAATCAATGAAACAGACACATT
TGGTCCTGGAGATGATGATGAAATCCAGTTTGACGATATTGGAGATGATGATGAAG
ACATTGATGATATCTAAATTGAACCAAGTGTTTTTACATGACAAGTTCTCTGAGGA
TGGTTCTACAGTTGGGATTTTGGCCATCATCAAC

~~Sequence ID - 638~~ SEQ ID NO: 239

nt:

545

TTTGAAGGCAAAGAGGGATTAATCTGTGCTGGCATCATGTAAGGAGACTTGATAGA
TAAGAAAAAGCTTTACCTAAGTTTTGAAGAATAGGTTTTTCATAATGGAAAATTTA
AGGGAAAAATCTCCAAAAAGTGCTACTCAAGTTTTATCCATTTGTATTTCCAACA
CAGCCTAGGACAGTACCTGCACATAGTAGGTGATTAATAAAAATTTAGAAAGCATT
AATACTAAAGAGGAAAAATAGCAATGGCAAGAAAACACATGTAGGGAACACATGTA
GCCAAAAAATAATATATAATCAGAGAAATAATAGGACTTCTGGAAAAAAAAGATGA
GATCAGATTGGTTAGGATCTTTACTAACATGACAAGAGCATGAATTTTTTTTCTGT
AGATAATAAGTATGAAAGAATTTTAGCTTAAAAATTAGCATAATTTGGATCCACAT
ATGCAAATCAATGAATGTAATTCATAATATAAACAGAACTAAACACAAAAACCACG
TGATTATCTCAATAGACACAGAAAAGGCCTTCAAAAAAATT

~~Sequence ID - 639~~ SEQ ID NO: 240

nt:

624

GACACACGAGCATATTTACCTCCGCTACCATAATCATCGCTATCCCCACCGGCGT
CAAAGTATTTAGCTGACTCGCCACACTCCACGGAAGCAATATGAAATGATCTGCTG
CAGTGCTCTGAGCCCTAGGATTCATCTTTCTTTTCACCGTAGGTGGCCTGACTGGC
ATTGTATTAGCAAATCATCACTAGACATCGTACTACACGACACGTACTACGTTGT
AGCCCACTTCCACTATGTCTTATCAATAGGAGCTGTATTTGCCATCATAGGAGGCT
TCATTCACTGATTTCCCCTATTCTCAGGCTACACCCTAGACCAAACCTACGCCAAA
ATCCATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCACAACACTT
TCTCGGCCTATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACACCA
CATGAAACATCCTATCATCTGTAGGCTCATTCATTTCTCTAACAGCAGTAATATTA
ATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAAGCGAAAAGTCCTAATAGTAGA
AGAACCCTCCATAAACCTGGAGTGACTATATGGATGCCCCCACCCTACCACACAT
TCGAAGAA

~~Sequence ID - 641~~ SEQ ID NO: 241

CAAGATGACAAAGAAAAGAAGGAACAATGGTCGTGCCAAAAAGGGCCGCGGCCACG
TGCAGCCTATTCGCTGCACTAACTGTGCCCCGATGCGTGCCCAAGGACAAGGCCATT
AAGAAATTCGTCATTCGAAACATAGTGGAGGCCGAGCAGTCAGGGACATTTCTGA
AGCGAGCGTCTTCGATGCCTATGTGCTTCCCAAGCTGTATGTGAAGCTACATTACT
GTGTGAGTTGTGCAATTCACAGCAAAGTAGTCAGGAATCGATCTCGTGAAGCCCGC
AAGGACCGAACACCCCCACCCCGATTTAGACCTGCGGGTGCTGCCCCACGTCCCCC

Marked-Up Copy

ACCAAAGCCCATGTAAGGAGCTGAGTTCTTAAAGACTGAAGACAGGCTATTCTCTG
GAGAAAAATAAAATGGAAATTGTACTTAA

~~Sequence ID 642~~ SEQ ID NO: 242

TGCTTGGCCCTCTACCTCCTGCCCTCTTCTGTTCATCTCCCAACCACTGCACTCT
TGATTTTTTATACCACACAGAAGGTAAGAAAATTCTAGGAACCCTAAGGATCAATCC
TCTCCATTTTCACTCAAATGCCTGGGGCCCAGCTCTGCAATGACTGACTCCAGGGC
CTCTTTTCTCACTGCCAGCATAGAAGTCAGGGGAGCCAGCTGGGGCCCTGCGGTCAG
GAAGGTTCTCATTTTTTGGAGCATTCCCTGAGCCCAGATCATAGGAGCAGCTGTCCC
TGGTGGGACACAGGAGTCATGACTCCTACCCTCCACCCTCCACACCCACCAGGCAT
TTAGCAGTCTGTCCTATGCAAGACAGATGAATTCTCAGCCAGGATACCTCAAGGCA
GGCAAAGGTGAGTGGAGGGAAAATTCACAAACATTCAGGGTGTGTGGTGCTGGCAT
CACCATGGCCAAATCCAAGAGGTCTTCCTGGAAGAGGGCCCAAACCTGGAACCAAAA
GAATGCTGTCAGCAGTTGGAATAGAGCTGTGAATT

~~Sequence ID 643~~ SEQ ID NO: 243

CTTTCCAAGAGGAATCCTCGGCAGATAAACTGGACTGTCCTCTACAGAAGGAAGCA
CAAAAAGGGACAGTCGGAAGAAATTCAAAAGAAAAGAACCCGCCGAGCAGTCAAAT
TCCAGAGGGCCATTACTGGTGCATCTCTTGCTGATATAATGGCCAAGAGGAATCAG
AAACCTGAAGTTAGAAAGGCTCAACGAGAACAAGCTATCAGGGCTGCTAAGGAAGC
AAAAAAGGCTAAGCAAGCATCTAAAAAGACTGCAATGGCTGCTGCTAAGGCACCTA
CAAAGGCAGCACCTAAGCAAAAGATTGTGAAGCCTGTGAAAGTTTCAGCTCCCCGA
GTTGGTGGAACGCTAAACTGGCAGATTAGATTTTTTAAATAAAGATTGGATTATA
ACTCT

~~Sequence ID 644~~ SEQ ID NO: 244

CTTTGATAGAGAAGAAAATTCTCCTAGGATACAAGAGCCTCAACATTTTAAAGATT
TTCTGCATCTCAAAGCGTAGGCTCCTTGCTGGGCAAGGTGAGCCTCTGTGAGTCC
TCATAGGACCGAGCAAATCTGATTACCCCCAGAAAATCCAATATCGAAGCTGAGCT
TTGGCCTGAGCGGGTTCCATTTTCTCCCCAGATCCTATTTAGGAAGTGTCTCCTGA
CAACCTCCAAAAGGTGCTAACATGCAACGTTCTGAAGGGTTATTGCTCAAAAACAA
GATTTTCTTGTGGTCAAGACTCTGCGAGCCTCGAACACGATGAATCCGCTCGAAT
GGGCTTGGGCTTTGCCCCGGGTGGCGCACGCTCACACGCTGGAAGCACAGCTTTGAC
GATCTCCACACACGCACAGGCACACACGCCACAGATGATGCCGGCTCATTCTCAGG
GGGTGTCTAAGTTCTGCTTTAAATATTTACCCCCTAATTGTACAAACAATAGGGGC

ATGAGCCTGGTACTCGATAAATGGGGACTTNCTTAAAA

~~Sequence ID - 645~~ SEQ ID NO: 245

nt:

649

CTACAGCCTGGGCAGCGCGCTGCGCCCCAGCACCAGCCGCAGCCTCTACGCCTCGT
CCCCGGGCGGGCGTGTATGCCACGCGCTCCTCTGCCGTGCGCCTGCGGAGCAGCGTG
CCCCGGGGTGCGGCTCCTGCAGGACTCGGTGGACTTCTCGCTGGCCGACGCCATCAA
CACCGAGTTCAAGAACACCCGCACCAACGAGAAGGTGGAGCTGCAGGAGCTGAATG
ACCGCTTCGCCAACTACATCGACAAGGTGCGCTTCCTGGAGCAGCAGAATAAGATC
CTGCTGGCCGAGCTCGAGCAGCTCAAGGGCCAAGGCAAGTCGCGCCTGGGGGACCT
CTACGAGGAGGAGATGCGGGAGCTGCGCCGGCAGGTGGACCAGCTAACCAACGACA
AAGCCCGCGTCGAGGTGGAGCGCGACAACCTGGCCGAGGACATCATGCGCCTCCGG
GAGAAATTGCAGGAGGAGATGCTTCAGAGAGAGGAAGCCGAAAACACCCTGCAATC
TTTCAGACAGGAAATCCAGGAGCTGCAGGCTCAGATTCAGGAACAGCATGTCCAAA
TCGATGTGGATGTTTCCAAGCCTGACCTCACGGCTGCCTTGCGTGACGTACGTANC
AATATGAAAGTGTGGCTGCCAAAAACCTTGCAG

~~Sequence ID - 646~~ SEQ ID NO: 246

nt:

600

GAGATGTCTCGCTCCGTGGCCTTAGCTGTGCTCGCGCTACTCTCTCTTTCTGGCCT
GGAGGCTATCCAGCGTACTCCAAAGATTCAGGTTTACTCACGTCATCCAGCAGAGA
ATGGAAAGTCAAATTTCTGAATTGCTATGTGTCTGGGTTTCATCCATCCGACATT
GAAGTTGACTTACTGAAGAATGGAGAGAGAATTGAAAAAGTGGAGCATTTCAGACTT
GTCTTTTCAGCAAGGACTGGTCTTTCTATCTCTTGTACTACACTGAATTCACCCCCA
CTGAAAAAGATGAGTATGCCTGCCGTGTGAACCATGTGACTTTGTACAGCCCAAG
ATAGTTAAGTGGGATCGAGACATGTAAGCAGCATCATGGAGGTTTGAAGATGCCGC
ATTTGGATTGGATGAATTCCAAATTCTGCTTGCTTGCTTTTTAATATTGATATGCT
TATACACTTACACTTTATGCACAAAATGTAGGGTTATAATAATGTTAACATGGACA
TGATCTTCTTTATAATTCTACTTTGAGTGCTGTCTCCATGTTTGATGTATCTGAGC
AGGGTGCTCCACAGGTAGCTCTAGGAGGGCTGGCAACTTA

~~Sequence ID - 647~~ SEQ ID NO: 247

CGAATGTGCAGGTTTGTACATAGGTATATATATGCCATGATGGAAATATTTATTT
TTTTAAGCGTAATTTTGCCAAATAATAAAAACAGAAGGAAATTGAGATTAGAGGGA
GGTGTTTTAAAGAGAGGTTATAGAGTAGAAGATTTGATGCTGGAGAGGTTAAGGTGC
AATAAGAATTTAGGGAGAAATGTTGTTTATTATTGGAGGGTAAATGATGTGGTGCC

Marked-Up Copy

TGAGGTCTGTACGTTACCTCTTAACAATTTCTGTCCTTCAGATGGAACTCTTTAA
CTTCTCGTAAAAGTCATATACCTATATAATAAAGCTACTGATTTCCAAAAA

~~Sequence ID 648~~SEQ ID NO: 248

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

~~Sequence ID 649~~SEQ ID NO: 249

nt:

425

CAAAAAACGAAGAAAAGTGACGACAGTCTGAGGGACTTATGGGAGATCATCAAGT
GAACCACTATATGTGTAATGTAAGTCTTGGAATGAGAAGAGAGAAGGAGAAGGAGG
AGAGAGCTTATTTGTAGAAATAATGGCTGAAAACATCCCAAACCTTCTTTTTTTT
AGGAAAGAAATAGGCATACAAGTTCAAGAACTCAAGGAACCTCAGAGAGGACAAT
TCTAAAGACACCCCTCTAACATACATTATAATCAAATTGTCAAAGTAAAATACA
AAGAGAATCTTTTAAATTGACAAGAGAAAAGCAGCTGGTCACGTTCAAGGGAGTTC
TATAAGAATTTTCAGCAGATTTCTCAGCAGAAACCTTGCAGGCCAACAGGCAGTGGG
ATGATACATTCAAAGTGCAAAAAAAAAAAAAAAAAA

~~Sequence ID 650~~SEQ ID NO: 250

CGAGAGTTTACCAGTNGCCTAATAATGCAATAAAAAATGCTTTGAGATAGCTAACN
GCCCATAAAACAACTCAAATTGCTTATAAAGTTTCTTCCCATGTTCCCATTTGAT
GAAAAGTCTTACATCACATATAACTGGGAAGCAGGGTCCCTCCTCAATTTTCAGA
CATTTTGAAAGGATGACAGTTCTGTTTGTAGATGAGTAAACCTCTATATTCATAA
GTTCTAAAATCCTTCATTATGAGGGATTCAAAGTATTTATAAAAACACTGCCCTCT
AAAAATTTCTCAGATCTGAAGTATGGNCTTGGNCTGAATATACAGTGTTATCCT
ATGTTTAAAAGGGTGATCCAGACATGAGACGCAACTAGTTGGTGCATAAGAAGGCC
CCACTTGGCTATTTCATATCTACCTACAATTGACCAAAAAAATTTTTTAGGCCAG
CAATTATTATTTAGCTTCGCTCTTCTAGTGCAAGAACTGCAGGCTGGATCAGTA
GTTCAACAGCTAAACAGTCATAAAATAGTCATTGGCATGTTAAATTTCTTTCAATG
CTTCAAAGATAAATTCCAATTCTATTTACTTATTCATTGNGACNGNATTACTAAAC
AGGTAAGGATGGGAATA

~~Sequence ID 651~~SEQ ID NO: 251

nt:

251

CTTTGGGAGGCCGAGGCGGGCGGATCACTTGAGGTCAGGGGTTGAGACCAGTCTG
GCCAACATGGTGAAACCCCACTCTACTAAAAATACAAAAGTTAGCCAAGTGTGGT
GGCAAGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATCACTTTGA

~~Sequence ID 652~~SEQ ID NO: 252

~~Sequence ID 653~~ SEQ ID NO: 253

~~Sequence ID 654~~SEQ ID NO: 254

~~Sequence ID 655~~SEQ ID NO: 255

[illegible]

~~Sequence ID 656~~ SEQ ID NO: 256

TAGAGGCCTGAATAGGTAGACAATGGCAGCAGCGTTTTTAATCACAGTCCTATTCA
TGCCCTAATTCGGGAGTGATGATTAAAGGACATTAGAGGGAGCACTTTGACATCTG
ATCCTTTGAACTGACGTCTGTGCAGGCTGCACTCCATAGAGCTCACTTGGCCAAAC
TGATTTTCCTTAAATAAAGTGCTGTGATTTCCAATGTAGGAAATATTACATTAGAGC
CTATTGAAATGATTAGGAATTGAGGAGCTTTTCTTTAGGTGGGAATGTGGTGTATG
CTGTATACTCACAAAAGTGAGATCATTAATATTGCATGTACTACTTTGAATATCAG
GGACCACAGAGAAATAGCATGAGAAACGCCTTCCTGCAGTCATGCACTTAAAATGA
ATATGAACAAAAATGTGGAACCTCTGCTGTCATAGCTCTCCG

~~Sequence ID 657~~ SEQ ID NO: 257

GGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAAGTTTTTTT
CTCTTTGAAAGATAGAGATTAATACAACCTTTAAAAAATATAGTCAATAGGTTACT
AAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAAGATTTTAAG
AGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAAAGGTTTCT
AAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAATGTATTTAAA
AGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAGAAGGGCAAT
GCTTTTAGATTAAAAATGAAGGTGACTTAAACAGCTTAAAGTTTA

~~Sequence ID 658~~ SEQ ID NO: 258

GACCTTTGAGAAAATTAATTTAAATCCTAGAACTTTGGGTGAACCGAAGAAATTTA
TAATATTTGTTTAGTTAATAACAGATAAAAAGGAAAGATTCAAGCCTATTGGATGA
GAATTTGTACATTATTTTAGAGCTAATAATAATGGTTTTTCAGTTTAGTGAGGATTT
AAAAAATGTTTTTGAATCAAACCTTTTTTTCTTTATAATCCTTTTTTAACATACTCAG
GAAATAAGGTATTATGAAATCCACACACTGTTACCTCCTTAAAGTATGAGGATACT
TCCCAGTGTGTTGGTCCACTAGTGGCTGATTATTTTGTGTTGTGGATTATTTGTAATT
TTCTTTTTTAATTCTTCCTTAAAGAGCATGGCATTGAGGATCACAGACCTATATTTG
AATCCTGTCATTTACTAGCGTTTTGACCTTGAACAATTATGCTCAGAGTCTCAGTT
TTTTCTTGTAAGTGATGATGATACTACTTAACTCACAGGGTTGTAGTGAAGATCA
AATGAGATCATGTCTGTANAACACCCTGCCCGGCACTCAATAAGTATTAATAGGAA
CCCATATACCTC

~~Sequence ID 660~~ SEQ ID NO: 259

TGTTTTTATTTTTTAAAAGGTATAAACACCAAAAAAAAAAATTAACATTGTATGAAG
ATGGAAAATAAGAAGATGCACTTTCTGTAACCTTGTCTAAGGATTTAAATTACTAA
CTTATGAACTCCAATTTGAATTGAACTTAACTATCGGCTTCTTACTGGTAAATTT

Marked-Up Copy

ATATGGTTTATTTTAAATGCGTACATATTGACCAATGGCCTCTGAAAAAGCACATT
TTAGATACTGAAATTGAAGGAAAGAAAATGCATCTTCAAACATTTTTTGGAAATCTC
ACCACATATACTTTGTTANATTTGTGTATTGTAGGGTGTTTGTGTTTTGTATTTTTGT
ATTGTATATGAACTTTTTTTAAATGTGACAGTTAAACACATCTTTAAAAGCATAGT
CACAGACAAAAGCATACAGTATAAAAATTTCTTGAAAACCTCTACAATATTATAT
TTGGAGGCAGCTTCAGACTGTTTTATTGG

~~Sequence ID - 661~~ SEQ ID NO: 260

CTCTGGCACACATTAGTTCCTCTTATATTACATTGATATAAGCAAGTCATATGGAT
TTATCTGAGTGTAAGGAGAGCTGGAAAAAATAGTTTCTAGCAGGTCAGCCACCTCC
CAGTGAGGGCTGCATACCATAGAAGGGGAGAATGAATTTTGGGAAAACAGGTAATT
ATCTCTGTCACAGAAGGGGATGAAAAGTATGGTAGTTACNCAAGTTANACATCTGT
ATGGAAAATACCACTTGGTTCTACAAATGNNG

~~Sequence ID - 663~~ SEQ ID NO: 261

nt:

627

GCCTCCCGGGTTCAGGGATTTCTCCTGCCTCAGCCTCCTGAGTGGCTGCATTGCAG
GCACCTGCCACCACGCCTTGCAAATTTTTGTGTTTTTAGTGGAGATGGGGTTTTGC
CATGTTGGCCAGGCTGGTCTCGGACTCCTGACCTCAGGTGATCCGCCCCGCTCAGC
CTCCCAGAGGGCTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCCCAAGTTTTGC
ATCTTTTAATGCCCTCTGAACAAATACATAGAGAAAACCTCTCAGAACAATTAAAAC
CTGCAGAGCAACAGTGTCTCCATGTCTTAGGTTTCAAGTTTGCCTCTAAAATTCT
AATCCATATTTTTCTACTTCTCAGATAATTTATGTGTGTGTACTCTTCCTAGACGT
ACAAGAGACTTTTTAATGCTAAATATTTGTCAGTGCTTAACAAAACTCAATTTCA
CATTACTCATATTGTTTTTGTGTTTTAATTGAATGTGAATTAAATTTTTATTAGTTAT
TTGATTTGGAATGTTATGTATGCCATTAACACTATTAGGGGAATCTCTAGCATTTCT
TGTATTTTTTAAAGAATTTGATTCTTTTGTANATTCTGCCTGTGTGGCATTTTAAAC
ATGTGTGACAT

~~Sequence ID - 665~~ SEQ ID NO: 262

nt:

345

ACCGGCGACATGGCCAAACGTACCAAGAAAGTCGGGATCGTCGGTAAATACGGGAC
CCGCTATGGGGCCTCCCTCCGAAAATGGTGAAGAAAATTGAAATCAGCCAGCACG
CCAAGTACACTTGCTCTTTCTGTGGCAAACCAAGATGAAGAGACGAGCTGTGGGG
ATCTGGCACTGTGGTTCCCTGCATGAAGACAGTGGCTGGCGGTGCCTGGACGTACAA
TACCACTTCCGCTGTCACGGTAAAGTCCGCCATCAGAAGACTGAAGGAGTTGAAAG

ACCAGTAGACGCTCCTCTACTCTTTGAGACATCACTGGCCTATAATAAATGGGTTA
ATTTATGTA

~~Sequence ID - 666~~ SEQ ID NO: 263

nt:

252

ATAATTCAGAACTTCTTCATATGCTCGAGTCTCCAGAGTCACTCCGTTCTAAGGTT
GATGAAGCTGTAGCTGTACTACAAGCCCACCAAGCTAAAGAGGCTGCCCAGAAAGC
AGTTAACAGTGCCACCGGTGTTCCAAGTGTAAAATTGATCAGGGACCATGAAAA
GAACTTGTGCTTCACCGAAGAAAAATATCTAAACATCGAAAACTTAAATATTAT
GGAAAAAAACATTGCAAAATATAAAAT

~~Sequence ID - 669~~ SEQ ID NO: 264

TTACTTTTAACCAGNGAAATTGACCTGCCCCGTGAANAGGCGGGCNTGACACAGCAA
GACGAGAAGACCCTATGGAGCTTTAATTTATTAATGCAAACGGTACCTAACAAACC
CACAGGTCCTAAACTACCAAACCTGCATTAAAAATTTTCGGTTGGGGCGACCTCGGA
GCAGAACCCAACCTCCGAGCAGTACATGCTAAGACTTCACCAGTCAAAGCGAACTA
CTATACTCAATTGATCCAATAACTTGACCAACGGAACAAGTTACCCTAGGGATAAC
AGCGCAATCCTATT

~~Sequence ID - 670~~ SEQ ID NO: 265

GGCTGATTCCCTGAGCTATAAAAGCATAATTGCTTTTATATTTTGGATCATTTTTTAC
TGGGGGCGGACTTGGGGGGGGTTGCATACAAAGATAACATATATATCCAACTTTCT
GAAATGAAATGTTTTTAGATTACTTTTTCAACTGTAAATAATGTACATTTAATGTC
ACAAGAAAAAATGTCTTCTGCAAATTTTCTAGTATAACAGAAATTTTTGTAGATG
AAAAAATCATTATGTTTAGAGGTCTAATGCTATGTTTTTCATATTACAGAGTGAAT
TTGTATTTAAACAAAAATTTAAATTTTGGAAATCCTCTAAACATTTTTGTATCTTTA
ATTGTTTATTATTAAATAAATCATATAAAAAT

~~Sequence ID - 671~~ SEQ ID NO: 266

CAGGAAGTCACCTGGGATTGGCTGCCTCACCCACTCACAGTGCCATCCCTGCCCCA
GGCCTCCCAGTGGCAATTCCAAACCTGGGTCCCTCCCTGAGCTCTCTGCCTTCTGC
TCTGTCTTTAATGCTACCAATGGGTATTGGGGATCGAGGGGTGATGTGTGGGTAC
CTGAAAGAACTACACCCTACCTCCACCACCTTACCCTCACCTGGAGAGCAGTTAT
TTCAGAACCATTTCTACCTGGCATTTTATCTTATTTAGCTGACAGACCACCTCCACA
GTACATCCACCCTAACTCTATAAATGTTGATGGTAATACAGCATTATCTATACCA
ATAACCCTTCAGCACTA

~~Sequence ID 672~~SEQ ID NO: 267

CAGGAAGTCACCTGGGATTGGCTGCCTCACCCACTCACAGTGCCATCCCTGCCCCA
GGCCTCCCAGTGGCAATTCCAAACCTGGGTCCCTCCCTGAGCTCTCTGCCTTCTGC
TCTGTCTTTAATGCTACCAATGGGTATTGGGGATCGAGGGGTGATGTGTGGGTAC
CTGAAAGAACTACACCCTACCTCCACCACCTTACCCTCACCTGGAGAGCAGTTAT
TTCANAACCATTTCTACCTGGCATTATTTATCTTATTTAGCTGACAGACCACCTCCACA
GTACATCCACCCTAACTCTATAAATGTTGATGGTAATACAGCATTATCTATCACCA
ATAACCCCTTCAGCACTAGATCCCTATCAGTCCAATGGAAATGTTGGATTANAACCA
GGCATTGTTTCAATANACTCTCGCTCTGTGAACACACATGG

~~Sequence ID 673~~SEQ ID NO: 268

GGGTTTTCTTTTCGGAAGCGCGCCTTGTGTTGGTACCCGGAATTCGCGGCCGCGTC
GACTGCTAAACAGAATACTGCTATTTTGAGAGAGTCAAGACTCTTTCTTAAGGGCC
AAGAAAGCCACNTGNNCCCTNGGNCTAATCTGGCTGAGTAGTCAGTTATAAAAGCC
NTAATNGCTTNNTNTTTGGNNTCNTTTTTNNCNGGGGNCGGNCTTGGGGGGGGTTG
CNTCCAAAGATANCATNTNTTTCCAACCTTNTNAANNNAANNNGTTTTAAAATCCCT
TTTCCNCCNGAAAANANNGCCCTTTAAGNGCCNCAAAAAAAAAANNGTNTTCTGCAN
NTTTTCTANTATNACAAANNTTTTNGTAGAANAAAAATTTTTTTTTAGNGGCTACC
CTTTNTTTNTTANNCANNGGAGTTTNTTTTTTACAAAAAAAAANATTGGGNCCCCCT
CCACAACCTTGGGTCTNTAATNGGGGGGTTTTTAAATAAANCNTNTNTAAATCCCC
CNNNNNNNNNCNNNNNNNNNCCNNNNNNNNNNNNNNNCCNNNNAAAAAATTTTTNC
TCCCCCNCCCTTTTTCTTCCTGCCGGCCCCAATTTAAGCCCGGGCGCTTGGGGCAA
ATCCCCCTTTAGNGGGGGGGTTTTANAAAAACNGGGGCGGGGNTTTAAACCNCGG
GGNNNGGGGAA

~~Sequence ID 674~~SEQ ID NO: 269

ACCTCTAGCATCACCAGTATTAGAGGCACCGCTGCCAGTGACACATG-TTTAAC
GGCCGCGGTACCCTAACCGTGCAAAGGTAGCATAATCACTTGTTCTTAATTAGGG
ACCTGTNTGAATGGCTCCACNAGGGTTCACCTGTCTCTTACTTTTAACCAGTGAAA
TTGACCTGCC

~~Sequence ID 675~~SEQ ID NO: 270

nt:

591

GTATAGAAAATAATGTCCCCAGNGCATAGAAAAAATGAGTCTCTGGGCCAGTGAAT
ACAAAACATCATGTGAGAATCATTGGAAGATATACAGAGTTCGTATTTTCAGCTTT

Marked-Up Copy

GTTTATCCTTCCTGTTAAGAGCCTCTGAGTTTTTAGTTTTAAAAGGATGAAAAGCT
TATGCAACATGCTCAGCAGGAGCTTCATCAACGATATATGTCAGATCTAAAGGTAT
ATTTTCATTCTGTAATTATGTTACATAAAAGCAATGTAAATCAGAATAAATATGTT
AGACCAGAATAAAATTAATTATATTCTGGTCTTCAAAGGACACACAGAACAGATAT
CAGCAGAATCACTTAATACTTCATAGAACAAAAATCACTCAAAACCTGTTTATAAC
CAAAGAATTCATGAAAAAGAAAGCCTTTGCCATTTGTCTTAGAAAGTTATTTTTTA
AAAAAAAATCATACTTACTATTAGTATCTATGGAAGTATATGTAACAATTTTTATG
TAAAGGTCATCTTTCTGTGATAGTGAAAAAATATGTCTTTACTAAGTTGAAATGAA
TACTTTCTGNCTTTGCTAATGGATAGTTATT

~~Sequence ID 676~~SEQ ID NO: 271

CTCAATTCTACTAAAAAGCCCCCAAGAAAAGCGAATGAGAAAACAGAGTCATCCT
CTGCACAGCAAGTAGCAGTGTACGCCTTAGCGCTTCCAGCTCCAGCTCAGATTCC
AGCTCCTCCTCTTCCTCGTCGTCTTCAGACACCAGTGATTCAGACTCAGGCTA
AGGGGTCAGGCCAGATGGGGCAGGAAGGCTNCGCAGGACCGGACCCCTAGACCACC
CTGCCCCACCTGCCCCCTTCCCCCTTTGCTGTGACACTTCTTCATCTCACCCCCCCC
TGCCCCCCTCTAGGAGAGCTGGCTCTGCAGTGGGGGAGGGATGCAGGGA

~~Sequence ID 679~~SEQ ID NO: 272

GNANCNTTTCCTNTCGNAAANCGCGCCTTGTGTTGGTACCCGGGAATTCGCGGCCG
CGTCGACAAAAAAAAAAAAAAAAAAAAAAAAAANTNTAGACTCGANCAAGCTT
ATGCANGCNTGCGGCCGCAATTCGAGCTCGGCCGACTTGGCCAATTCGCCCTATAG
NGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGNGACTGGGAAAACC
CTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGT
AATANCGAANAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGG
CGAANGGAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTA
AATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGC AAAATCCCTTATAAATCAA
AAGAATAGACCGAGATAGGGTTGAGNGTTGTTCCAGTTTGGAACAANAGTCCACTN
TTAAAGAACGNGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGG
CCCCTACGTGAACCATCNCCCTAATCAAGTTTTTTGGGGTCGAGGNGCCGTAAAG
CACTAAATCGGAACCCTAAAGGGAGCCCCGATTTAAAGCTTGACGGGGAAAGCCC
GGCGAACGTGGCGAAA

~~Sequence ID 682~~SEQ ID NO: 273

CACCTGCAGTCCAAGTACATCGGCACGGGCCACGCCGACACCACCAAGTGGGAGTG
GCTGGTGAACCAACACCGCGACTCGTACTGCTCCTACATGGGCCACTTCGACCTTC

Marked-Up Copy

TCAACTACTTCGCCATTGCGGAGAATGAGAGCAAAGCGCGAGTCCGCTTCAACTTG
ATGGAAAAGATGCTTCAGCCTTGTGGACCGCCAGCCGACAAGCCCGAGGAAAAC TG
AAACTTTGCTTAACNACCGAATGGNGGGGANCTTTTCCAACGNTTTT

~~Sequence ID 683~~SEQ ID NO: 274

TTGGTTTTCATACTGNTGGGGNTTGAATGNTCCCTNCAACACTNATGTTGANACTTA
ATCCCTAATGNGGCAATACTGAAAGGTGGGGCCTTTGAGATGTGATTGGATCGTAA
GGCTGTGCCTTCATTCATGGGTAAATGGATTAATGGGTATCACAGGAATGGGACT
GGTGGCTTTATAAGAAGAGGAAAAGAGAACTGAGCTTGCATGCCC

~~Sequence ID 684~~SEQ ID NO: 275

nt:

545

GTGGAAGNGACATCGTCTTTAAACCCTGCGTGGCAATCCCTGACGCACCGCCGTGA
TGCCCANNGAAGACAGGGCGACCTGGAAGTCCAAC TACTTCCTTAAGATCATCCAA
CTATTGGATGATTATCCGAAATGTTTCATTGTGGGAGCAGACAATGTGGGCTCCAA
GCAGATGCAGCAGATCCGCATGTCCCTTCNCGGGAAGGCTGTGGTGCTGATGGGCA
AGAACACCATGATGCGCAAGGCCATCCGAGGGCACCTGGAAAACAACCCAGCTCTG
GAGAAACTGCTGCCTCATATCCGGGGGAATGTGGGCTTTGTGTTCACCAAGGAGGA
CCTCACTGANATCAGGGACATGTTGCTGGCCAATAAGGTGCCAGCTGCTGCCCCGTG
CTGGTGCCATTGCCCCATGTGAAGTCACTGTGCCAGCCCAGAACACTGGTCTCGGG
CCCGATAAGACCTCCTTTTTTCCAGGCTTTAGGTATCACCAC TAAATCTCCAGGGG
CACCATTGAAATCCTGAGTGATGTGCACTGATCAAGACTGG

~~Sequence ID 685~~SEQ ID NO: 276

GGAAAGGGCCATTTTATTGCCTAAAACCACCTGGNTTTTNAGGTAACAGTTCCAAC
ATGTCCTTTTTTTGAATAGCTGTTCTAATTATTATATATTCAGCTGATTAATAGGAG
TACTTGATAGGTGGACTGTGTCAGGTAGCCTCAGGCAATCCTACTTCAACAAGCTG
TCAGGGAGCCATGCCATGCTTCTTTATGACATAGGTGAATTTGATAGGCTCACTAG
CAGAACATGGGATCACAAGGTGGAACNTTCCNTTT

~~Sequence ID 686~~SEQ ID NO: 277

GACCCCTTCCTTACACCTTATACAAAAAACTGAAACTGGACCCCTTCCTTACACC
TTATACAAAAATTA ACTCAATTTTATTATGTTGTATTAAATTAAGTTGGGTTTAAT
TAAGATGGATTAAAGACTTAATTATAAGACCTAAAACCATAAAAACCCTAGAAAGAA
AACCTAGGCCATACCATTCAGGACACGGGTATGGGCAAAGACTTCATAACTAAAAC
ACCAAAAGCAATGGCAACGAAGTCCAATAGACAAATTGGACCTGATTAAACTAAA

Marked-Up Copy

GAGCTTCAGCACAGCAGAAGAGACTATCGTCAGAGTGAACAGGCAACCCACAGAAT
GGAAGAAAATTCTTGCAATCTATCCATCTGACAAGGGGCTAATATCCAAAATCTAC
AAAGAACTTAAACAAATTTACAAGGAAAAACACAAACAACCCCATCAAAAAGTGGG
CTAAGGATGTGAACAGACACTTCTCAAAAGAAAACATTTATGCAGCCAACAAACAT
GAAAAAAAGTTCATCATCACTGCTCATTAGAGACATGCAAATCAAAACCACAATGA
GATCCCATCCACACCAGTTAGAATGGCAATCATTTAAAAATGT

~~Sequence ID - 687~~ SEQ ID NO: 278

nt:

268

TTTATGTGTTTTTGCTTGGGGGGCGCTGGGCCTAGCCCAGAGTAGTGCTTGCTCCC
CCTGCCTTGTCCCACCAGGGAGGCAGCAGACTCAGGCCCTCCATGGTCCTCTTTGT
CATTTTGTGACATGCATTCTCCTTTTGTTCATCTTGTGGGGGGAGGGGATTAAC
CAAAGGCCACCCTGACTTTGTTTTTGTGGACACACAATAAAAGCCCCGTTTATTTG
TAAA

~~Sequence ID - 688~~ SEQ ID NO: 279

nt:

569

CTTTAGCCAGCCTGATCAGAAAAAACAAAAGAAGAGGAAAGACGTAGATTACCAA
CATCAAGAATGTGAGTTATGATATCACTACAGACTCTCCAGGTATTAAAAGCATAA
TTAGAGAATGATATGAGCAGCTATATGCAAATAAGTTCAACATTGGACAAATGGAC
AAATTTCTTGAAAGATAAATTATGAAATTTCAATCTGAAAGAACTACATGACCTTA
ATTGTCTTACATCTATTAAATAAGTGGAATTTGTAGTTTAGAACTTTCCCACAAA
GAAAACTCTAGGCCCAGATGGCATCAAAATAATATTCAGATGAATGAAATGGAGAA
AGGATAGCCTTTTCAACAAATGGTGGTGGACAATTGGATTTCCATATGCAAAAAA
ATAGAGATGGACGCAGAGGTGTGTGCTTAGGAGGCTGAGGTGAGAGGATTGTTTGA
GGCCAGCCTGGGCAACATAGCAAGACCCCATTTCAAAAACAAAAATAAAGAACTTG
TAGCCTTACCTTGTGCCATATTATGAAAATGTATCATAGGCTTAAATGTGAAACGT
AAAACAAA

~~Sequence ID - 689~~ SEQ ID NO: 280

CGCAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAAGTTT
TTTTCTCTTTGAAAGATAGAGATTAATACAACTACTTAAAAATATAGTCAATAGG
TTACTAAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAAGATT
TTAAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAAAAGG
TTTCTAAACATGACGGAGGTGAGATGAAGCTTCTTCATGGAGTAAAAATGTAT
TTAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACTAATAGAAGG

Marked-Up Copy

GCAATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTTAAA
AGTTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAAGAGATTAAACCGA
AGGTGATTAAAAGACCTTGAAATCCATGACGCAGGGAGAATTGC

~~Sequence ID 690~~ SEQ ID NO: 281

CGAAAAGCAAATATAACTTGCCACTAACCAAGATCACCTCTGCAAAAAGAAATGAA
AACAACTTTTGGCAGGATTCTGTTTCATCTGACAGAATTCAGAAGCAGGAAAAAAA
GCCTTTTAAAAATACCGAGAACATTAAAAATTCGCATTTGAAGAAATCAGCATTTTC
TAACTGAAGTGAGCCAAAAGGAAAATTATGCTGGGGCAAAGTTTAGTGATCCACCT
TCTCCTAGTGTTCTTCCAAAGCCTCCTAGTCACTGGATGGGAAGCACTGTTGAAAA
TTCCAACCAAAACAGGGAGCTGATGGCAGTACACTTAAAAACGCTCCTCAAAGTTC
AAACTTAGATTTTCAGATTT

~~Sequence ID 691~~ SEQ ID NO: 282

CCGGTCTCTACACAATATATAGAAATCTGGGCATGGTGGTGCCTGGCTGTAGTCTC
AGCTACCTAGTTGGGTGAGGTGGGAGAGTCGCTTGAGTCCTGGAGGTTGAGGCTGT
AGTGAGCCAGGGCTGCACCACTGCATTCCAGCCTGGGTAACAGAGTGAGACCCTGT
CTCAAAAAGAAAAAAAAAAAAAATTGCTAATTTTAACAAATCACAAAACCTGACTCAGGC
AAGTTGTCTGACTCAAAGCCCTTGAAAAACCATCAAAGACAGTAGAATGTAACT
GGTCATTTACGTAAAATAGTGTTCAATTAATTTTTGGTTCAATTTAGGATAATCAT
TTAAATGAGACTGTATTTGAGACTGTATACACATACATATACATGTTTACACACAT
ATACGTACAATATATGTACATTCTATCTAAAAGATCATAACATGTGTGTACATATAT
GTTTTTAAAAGTCAAACCTGACATATTAATGGAAACAGTGCTTACATCTCTGGTAGT
GATTTTCTATTAGCAGCAGCCCTACATATGCTGCGTCTCTGAACAGCATGTCAGTG
CCATGACTGTCTAAACATGCAAATATGACTGACAGACTCTTGAGACAGCTTTCACC
TTG

~~Sequence ID 692~~ SEQ ID NO: 283

AATTCGNGGCCGCGTCNNCCTANGAGGCACCAGGAAATCCCGCGGGGTGGCCCATG
CAGACCAGGCGCACGTGGCTCATGGGGCANAATTGCCAAGGACAGCTCACGACAGT
GCCACCTTCTCACCATTCCAGCCAAGGAGAGATGTGACGTTGGAACCTGCTCTGGCA
CTTCTGTCAAGCCTCCCCCGCCCCAATTGCCTTGAGATCTCTGCTCTTTGTCAGAG
ATTTGCAAAGACTCACGTTTTTGTGTTTCTCATCATTCATTGTGATACTAAGA
AACTAAGAAGCTTAATGAAAAGAAATAAAATGCCTATGTTGTTGTTCT

~~Sequence ID 693~~ SEQ ID NO: 284

Marked-Up Copy

CTAGAACCCATGACTCCTAGGTCTTATACTGCAACCACAGTATCAGCAAATAATCT
TTCATAAGGGGATTATTCTCTGATTAAACAGGAAATACAGGAATTTAATTTGTGAAC
ACGCTAGGTAGAAGCAGAAACCCAAATCCAAATCCAAATTTAAACATTTAAAATTC
ATTCTATAACTAAGATCTAACAGTCATTTTCTTCCCAGTAAGAAATAACCAAAGCA
TGCTAAAAATCACTGGACTAAATTGGTGTCAAACTGCCACATTGCCAGGCATGGG
GGGGTCATACTTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGAAAATTGCTTGA
GGCCAGGAGTTTCGAAACCAGCCTGGGCAACACAGTGAGACCCCATCTCCACAAAAA
AAAAAAATTAAAAAACAAAACAAAACATTAGCTGGGCATGGTGGTACACGCCTGTA
GTCCCAGCTACTCAGGAGCCTGAAGTGAGAGGATCACTGAAGCCCAGGAGGTAGAG
CTATGACTGTAGTGAGCTATGACTGTGCCACTACACTCCACCTGGGTGACAGGGGA
CTC

~~Sequence ID 694~~SEQ ID NO: 285

CGACTTCCATTTGTATTAATGGAATACTAAGTCCCTCTGTGATTTCTGAACCAAGC
TATTCCTAGGCCTGAGTTTTATTTTGTGACACAGAAATAAATTANAAGGCCAAGC
GTGGTGGCATGTGCCTGTAGTCCTAGTTGCTGAGGTAAGAGGATTGCTTGAGCCCA
GGAGTTCAAGGCTGCAGCAAGCTTTGATTGCGCCACTGCACTCCAGCCTTGGCGAC
AGACTAAGACGCTGTCTCAAAAAAAAAACAAAA

~~Sequence ID 696~~SEQ ID NO: 286

GGTTATCAATGAGATTAAGAGACAACCTAGAGTAAAAACAAAAGAAAAGAAAAGAAA
NGAAAACAACAGAAGCTCTATTAACCTGACCTCTAACCAATACAACAGGTAACTGA
TGTTCTCCATTCTGTATATAAAAATCCCAGTGGACACCCACAACACAGGCTTCAGG
CTTGTAGGACACTTTCTAGTTCATCTGAGCACTTTTGTCTCAGCAGTTGAGCTGT
ATACTTAGCAACATTTGGTGCTTCCAAACCCATTTGTGCCTGTAGCACTTACTATT
GAAATACATAATTTAATTAAATATTATATAAAGGAATGGAATACGAGTTGGACAAG
AAAAAGAGTTAAATCTGAAGGTTAGGTAAAAAGAGCAACTTCTTTTCTCTGTTTTG
CAGGTTGGCAAAATCATTTAAAAACAATTGGAAGTATTATATGTTCTGCATTAAGT
TGTCATTTTACTTAAAACTAGGCATCAAAGATGATGCATAATAAATTTAGTGTAT
GCAAGAATGACTGCTTGGGACCTCAATATATGAATTCTTAATCCAAGGAAAGTCCT
TGGCCTTACATTTAAAAGTCGGCAAATAAGTGTACGTTTCATT

~~Sequence ID 697~~SEQ ID NO: 287

GAACATTTAAAAATAATGCAAATAAGGCTGGGCGTGGGGGCTCACACCTGTAATCC
CAGCACTTTGGGAGGCCGAGGCAGGCAGATCACGAGGTCAGGAGATTGAGACCATC
CTGGCTAACACAGTGAAACCCTGTCTCTACTTAAAAAATAAAAAAATTAGCCAGGC

Marked-Up Copy

GTGGTGGTGGGCGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGGT
GTGAACCCGGGAGGCGGAGCTTGCANTGAGCTGAGATCGTGCCACTGCACTCCAGC
CTGAGCGACAGAGCGAGACTCTGTCT

~~Sequence ID 698~~ SEQ ID NO: 288

TCATTAGAATCCAAGCTTTGAAAATTTCTGATTAATGCTCATGTATTTCTTTATCT
TTGTTTTTCTTGTGAAGAAAGACTTTCACCACTGTCTGAGTGATGATGCTGTTGA
TAAGGATGATGTCGATGACTACTATATTGCATCTCTCAGGAACAGCTGATGGGAAG
GGAGGGGCTGCTGAGTTCCCTTGTTCTAGCTAGCAGCACGCTCCTCANAGAGGGGG
CCGAGTTACAGACAGCAGCCGCATTCTCATGCAAAATTAGTTTTAAACTGCTAGTG
TGGGCATCGGTACCTTTTGCCTGGGTGATACCGAAGAATTGTTGAGGATTTAGTAT
GCTCCGTAGAGACAGTTCAGCCAGTCATTTCTGCATTGGAGAGACTTCTCATACTT
TCTTTGAAGACTCATAGAAAGCTGGAT

~~Sequence ID 699~~ SEQ ID NO: 289

ATTAAGGTTTGTNCCCAACAAGAATAGATGTAATTAGAAAAAANTGNCTTCCTTAC
CTATTGCCTCTGATNTTTACTTGCTTAAATTTTTTTTATTGNAAATCCAGAAAAAG
NGGATTTAGAGAACAACACTAACTCCACCTAATCTATGACAGANATGTACAANAN
AGTACCTGTGAAAAATGTGAAAGNATNTGAAAAATGTAACCTTTGGCAGCCTGAGC
ATAGTCAACCAGAAAACTATCTGAATTAAATAATTGGTCCATAGGTACTATTTT
ATTTGGTCCATAAGGATTATTTTTTCAACTTTTTTTTCAAGTGTATTATTATGTCA
TTTCCCACGTAGGTTACTGATACCTGAAGACTTTTTTNCACCTTTAACCTTNCTCGT
TGAGGAGCTTTGTANTCTAATAAAAGAGAAATATAAGTAAATGTTAGATATATGGG
NGGATAATGGTAACTATGTGCTTAAAGAGGTATAAAAGAAGGGTAGGGAGCAGATA
AGACAAAGGAAGGGCTATATTATAANGAAGAATATTCCAAGTAGGGAAGAGAAAAA
GATATGTTATCCATATAATATTTTATGTGCAGTAGAGAACATGTTCTATAGAAANAG
ACAGAAGATG

~~Sequence ID 700~~ SEQ ID NO: 290

CTTGAGCCCAGGAATTCCAGCCTGGGCAATATAGTAAGACTCCGTCTCTACAAAAG
ATACAAAAATTAGCCAGATGTGGTGGTGCCTGTAGTTCCAGATACTGGAAAG
ACTGAGGCAGGAGGATTGCTTGAGCATGGGAAGTTGAGGCTGCAATGAGCTGTGAT
TACGCCACTACACTCCAGCCTGGGCAACAGAGTAAGATCTTGTCTCAAAAAAAAAA
TTGAATTCAGCTAAAAATAATAAAATTTTAAATAATTTTAAAAAGCCCTCAACAG
CTTTGTTTTTCTCTCCTTGCCAGCTTCTCTGCAGCCTATAGCCTGCAGGCTGGCTG
CTGCGAGCCAGGACAAGCGGTGGGAAATGCAATCACAGCGTGAAATCTCTGTGTTC

Marked-Up Copy

AGAGACACGCAGGAAGCAGGTGAACCATGAAGGGCCAACACATGCCCCCAGTTAGC
AGGGTGTAGAGACCGGGGCAGGGCTTTCTTCTTCCTTCTGGGTATATAATATCCAT
GTCCTGCCATTTGAAGCTGCAAGTGGCACACATGGATGCTGGACAGGCGCTCGCAC
TTTCTGGGCAGGGCANGGGGCTCAAAGGCAGGACAGCTGGGCAAAGCACCTTGCG
TGGGCCC

~~Sequence ID - 701~~ SEQ ID NO: 291

nt:

579

CTTTGGAGCTTCTGTCTGTGCTGTGGACCTCAATGCAGATGGCTTCTCAGATCTGC
TCGTGGGAGCACCCATGCAGAGCACCATCAGAGAGGAAGGAAGAGTGTTTGTGTAC
ATCAACTCTGGCTCGGGAGCAGTAATGAATGCAATGGAAACAAACCTCGTTGGAAG
TGACAAATATGCTGCAAGATTTGGGGAATCTATAGTTAATCTTGGCGACATTGACA
ATGATGGCTTTGAAGGTAATTAAATTTATCAAATTGGTGCTTGATTTCTGCTTTTA
AAATGGTTTATGGAAGAAAATATGATTAAAGTTTTGTATTGTTTTCTTCCTATAG
AAGATGGAGCCAGAATGGCATGCTAAGTTTTTTCTTTTCTTTAGTGTTATATATGA
CTTCTCCTCAATTGTCACCCATTGATCTTTACCACTGTTAATAATGGATGATATTC
AAAATACCTTATTTTCTAGTGATTCTAAGGCACCATTGATTAGAACTGCATTATTAT
TTATGTGTCCCTAAAAGCTACCTATTAAGCTGTTACACCCACCATTTTTCTGTAA
GAAAATCCTGATTTTCTAGAA

~~Sequence ID - 702~~ SEQ ID NO: 292

GTNNTCCTCTCGGAACGCGCCTTNTGTAGCCAGGTGCTACCAGACCNAATACACGG
TTGTTCCAGCTTGCGCATTACCCGATGGCGTAGATATCCGGATCGGAAGTCTGGCA
GGAATCATTAATGACAATACCCCCACGCGGAGCAACGTCCAGACCACACTGGGTTG
CCAGCTTATCGCGCGGACGGATACCGGTAGAGAAGACGATAAAGTCGACTTCCAGT
TCGCTGCCGTGGGCAAAACGCATGGTTTTACGCGCTTCAACACCTTCCTGCACAAT
CTCAAGGGTGTTTTTGTGCTGGTGTGAACGCGCACGCCATACTTTTCGATTTTTCGAC
GCAGCTGCTCGCCGCCCATCTGATCAAGCTGTTCTGCCATCAGCATAGGGGCAAAT
TCGATAAC
GTGGGTTTCAATACCTAAGTTTTTTCAGCGCGCCTGCGGCTTCCAGACCTAACAGGC
CGCAATTCGAGCTCGGCCGACTTGGCCAATTCGCCCTATAGTGAGTCGTATTACAA
TTCCTGCGCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAAC
TTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAAGAGGC
CCGCACCCGATCGCCCTTTCCAACAGTTGCGCACCTGAATGGCGAATGGAAATTGT
AAGCGTTAATATTTTGTAAATTCGCGT

Marked-Up Copy

~~Sequence ID 703~~ SEQ ID NO: 293

CTGCGCAGACCAGACTTCGCTCGTACTCGTGCGCCTCGCTTCGCTTTTCCTCCGCA
ACCATGTCTGACAAACCCGATATGGCTGAGATCGAGAAATTCGATAAGTCGAAACT
GAAGAAGACAGAGACGCAAGAGAAAAATCCACTGCCTTCCAAAGAAACGATTGAAC
AGGAGAAGCAAGCAGGCGAATCGTAATGAGGCGTGCGCCGCCAATATGCACTGTAC
ATTCCACAAGCATTGCCTTCTTATTTTACTTCTTTTAGCTGTTTAACTTTGTAAGA
TGCAAAGAGGTTGGATCAAGTTTAAATGACTGTGCTGCCCCTTTCACATCAAAGAA
CTACTGACAACGAAGGCCGCGCCTGCCTTTCCCATCTGTCTATCTATCTGGCTGGC
AGGGAAGGAAAGAACTTGCAATGTTGGTGAAGGAAGAAGTGGGGTGAAGAAGTGGG
GGTGGGACGACAGTGAAATCTAA

~~Sequence ID 704~~ SEQ ID NO: 294

CTTGTATTCAAGAACTACTGTAATGCATTAGTGGTCTGGCTTCATTTTGTATGATG
CCAGATCCTTAATTTACCCAGCACAAATCATTTTCAGTAGTTTCCTATGGCTCCTGCA
AAAATGCAAACAGAAACCACCACAGGAACAGCCCCTTGCTGCCTCCTGTTGCTGAG
GTAGTAGTCGCTAAAGAAAATTGAAGGCTCCTTACAATCTATATTTGAAAAGTAGA
ACTTCTGTAGAAACACACAGATCCCGATCTTAGAAGTTGTACAGGACAATCTGGTA
AAACTGACATAATTGTGATTTATTAACATGAATTAAAATGCCCAACCAGTGCTTCA
GTGTGACAGTATATTTAAAATAAAAAAGAAATTAAAGGTCATATACTGTACTACTT
TCACAAAGATCCACAGTTTTTGCAAAGACTTGTCATATGTACAATGCTATATATCA
AATGAGAAAAGCTGTAAGCAATTATATACGCAAAAGAAATGGCAGTA

~~Sequence ID 705~~ SEQ ID NO: 295

TTCCAGTCCTTTTCATTTAGTATAAAAGAAATACTGAACAAGCCAGTGGGATGGAAT
TGAAAGAACTAATCATGAGGACTCTGTCTTGACACAGGTCCTCAAAGCTAGCAGAG
ATACGCAGACATTGTGGCATCTGGGTAGAAGAATACTGTATTGTGTGTGCAGTGCA
CAGTGTGTGGTGTGTGCACACTCATTCCTTCTGCTCTTGGGCACAGGCAGTGGGTG
TAGAGGTAACCAAGTAGCTTTGAGAAGCTACATGTAGCTCACCAGTGGTTTTCTCTA
AGGAATCACAAAGGTAAACTACCCAACCACATGCCACGTAATATTTAGCCATTCA
GAGGAACTGTTTTCTCTTTATTTGCTTATATGTTAATATGGTTTTTTAAATTGGTA
ACTTTTATATAGTATGGTAACAGTATGTTAATACACACATACATATGCACACATGC
TTTGGGTCCTTCCATAATACTTTTATATTTGTAAATCAATGTTTTTGGAGCAATCC
CAAGTTTAAGGGAAATATTTTTGTAAA

~~Sequence ID 706~~ SEQ ID NO: 296

nt:

Marked-Up Copy

CAACCCCTCTCTCCTCAGCGCTTCTTCTTTCTTGGTTTGATCCTGACTGCTGTCATG
GCGTGCCCTCTGGAGAAGGCCCTGGATGTGATGGTGTCCACCTTCCACAAGTACTC
GGGCAAAGAGGGTGACAAGTTCAAGCTCAACAAGTCAGAACTAAAGGAGCTGCTGA
CCCGGGAGCTGCCCAGCTTCTTGGGGAAAAGGACAGATGAAGCTGCTTTCCAGAAG
CTGATGAGCAACTTGGACAGCAACAGGGACAACGAGGTGGACTTCCAAGAGTACTG
TGTCTTCCTGTCCTGCATCGCCATGATGTGTAACGAATTCCTTTGAAGGCTTCCCAG
ATAAGCAGCCCAGGAAGAAATGAAAACCTCCTCTGATGTGGTTGGGGGGTCTGCCAG
CTGGGGCCCTCCCTGTCGCCAGTGGGCACTTTTTTTTTTCCACCCTGGCTCCTTCA
ACACGTGCTTGATGCTGAGCAAAGTTCAATAAAGATTTTGGGAAGTTT

~~Sequence ID - 707~~ SEQ ID NO: 297

nt:

397

CGGATGTGGTGGCAGGCGCCTCTAGTCCCAGCTACTCGGCAGGCTGAGGTAGGAGA
ATGGCTTGAACCCAGGAGGTGGAGCTGACAGTGAGCCGAGATCGCGCCACTGCACT
CCAGCCTGGGCGGCAGAGCGAGACTCCATCTCAAAAAAAAAAAAAAAAAAATAGA
CTTTGAGACCAGCCTGACCAACATAGTGAAACCCGTCCTACTACTAAAAATACAAAA
TTACCCGGGCGTGGTGACGGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGAGACA
GGAGAATCACTTGAACCAGGGAGGCGGAGGTTGTAGTGAATGAAATCGTGCCCCCT
GCACTCCAGCCTGGGTAAACAAGAGCGAAACTCCGTCTCAAAAATAAATAAATAAAT
AAAAAT

~~Sequence ID - 708~~ SEQ ID NO: 298

nt:

293

CCAGCTTTTTATGGTGTTTAATCTAATACACTTAAGCTGCAGTCCCAAATTAGGG
GTCCTTCAGTCTTGGAGACTATAAGGGAGCCTCTGCACCCAGGGAAAATGTTACCC
TTTACAGGGGGGAAGGGTAAACCAGTAGGGAATACAGTACAATCCCAACCCTACTG
GGAGGGGCGGGAGGGAGGTGTTGCCGTCACTGTATTAAGTCGATGTTGGGAAACGT
TTTAACATCTGGAGCCTTTGTGGGTGGAAATATGTCTCCAGTTACAACCTCCGCAGT
GGATGTGAAGAAG

~~Sequence ID - 709~~ SEQ ID NO: 299

GGAAGCTACAATGATTTTGGGAATTACAACAATCAGTCTTCAAATTTTGGACCCAT
GAAGGGAGGAAATTTTGGAGGCAGAAGCTCTGGCCCCTATGGCGGTGGAGGCCAAT
ACTTTGCAAACACGAAACCAAGGTGGCTATGGCGGTTCAGCAGCAGCAGTAGC
TATGGCAGTGGCAGAAGATTTTAATTAGGAAACAAAGCTTANCAGGAGAGGAGAGC
CAGAGAAGTGACAGGGAAGCTACAGGTTACAACAGATTTGTGAACTCAGCCAAGCA

Marked-Up Copy

CAGTGGTGGCAGGGCCTAGCTGCTACAAAGAAGACATGTTTTAGACAAATACTCAT
GTGTATGGGCAAAAACTCGAGGACTGTATTTGTGACTAATTGTATAACAGGTTAT
TTTAGTTTCTGTTCTGTGGAAAGTGTAAGCATTCCAACAAAGGGGTTTTAATGTA
NATT

~~Sequence ID 710~~SEQ ID NO: 300

TGGATTCCCGTCGTAACCTTAAAGGGAACTTTCACAATGTCCGGAGCCCTTGATGT
CCTGCAAATGAAGGAGGAGGATGTCCTTAAGTTCCTTGCAGCAGGAACCCACTTAG
GTGGCACCAATCTTGACTTCCAGATGGAACAGTACATCTATAAAAGGAAAAGTGAT
GGCATCTATATCATAAATCTCAAGAGGACCTGGGAGAAGCTTCTGCTGGCAGCTCG
TGCAATTGTTGCCATTGAAAACCCTGCTGATGTCAGTGTTATATCCTCCAGGAATA
CTGGCCAGAGGGCTGTGCTGAAGTTTGCTGCTGCCACTGGAGCCACTCCAATTGCT
GGCCGCTTCACTCCTGGAACCTTCACTAACCAGATCCAGGCAGCCTTCCGGGAGCC
ACGGCTTCTTGTTGTTACTGACCCCAGGGCTGACCACCAGCCTCTCACGGAGGCAT
CTTATGTTAACCTACCTACCATTGCCCTGTGT

~~Sequence ID 711~~SEQ ID NO: 301

nt:

498

GTGGTACATATACACAAAGGAAAACCTATGTAGCCATTAAAAGAAAAGGAACTCCTA
TCATTTGTAACAACATAAATAAATCTGGAGGAGATTAGGCTAAGGTGAAATAAGCC
AGGCACAAAAAGACAACTACCATATGATCTTACTTATACGTGTGTGGAATCTAAAA
AGGTGGAATTTACAGAAGCAGAGAGTAGAATGGTGATTACCAGAGGCTGGGGAGTG
AGGGCAGGAGGTTGGAGAAATGTTGGTCAAAGGATACAAAGTTTCAGTTATACAGG
ATGAATAAGTTCAAGAGATCTATTGTACAACGTGGTGGCTATAGTTGATAACAATG
TATTGTGTTCTTGAAAAATGCTGAGAGAGTAGATTTTAAGTGTTCTCACCACAAAA
CATAAGTATGTGAGGTAATGCATGTGTTAATTANCTTAATTTAGACATTTTCATAAT
GTATTATACATATTTCAAACACGTTGTACATGAGAAAGATACACAATT

~~Sequence ID 713~~SEQ ID NO: 302

GCCCAGTCGACCCATGTTCTCCTTTCTACACCAGCATTAGACGCTGTCTTCACAGA
TTTGGAATCCTGGCTGCCATTTTTGCAGCTGCCATCCATGACGTTGATCATCCTG
GAGTCTCCAATCAGTTTCTCATCAACACAAATTCAGAACTTGCTTTGATGTATAAT
GATGAATCTGTGTTGGAAAATCATCACCTTGCTGTGGGTTTCAAACGCTGCAAGA
AGAACACTGTGACATCTTCATGAATCTCACCAAGAAGCAGCGTCAGACACTCAGGA
AGATGGTTATTGACATGGTGTAGCAACTGATATGTCTAAACATATGAGCCTGCTG
GCAGACCTGAAGACAATGGTAGAAACGAAGAAAGTTACAAGTTCAGGCGTTCTTCT

CCTAGACAACCTATAACCCGATCGCATTCAGGTCCTTCGCAACATGGTCACTGTGCAG
ACCTGAGCAACCCCAACCAAGTCCTTG

~~Sequence ID 714~~SEQ ID NO: 303

CTGTAACAGAGATTCCTTTTTTCAATAATCTTAATTCAAAAGCATTATTAGACTTG
AAAGGGTTTGATAATCTCCCAGTCCTTAGTAAAGATTGAGAGAGGCTGGAGCAGTT
TTCAGTTTTAAATGAGTCTGCAGTTAATATCAAATGTGAGTTTGGGACTGCCTGGC
AACATTTATATTTCTTATTCAGAACCCTTGATGAGACTATTTTTAAACATACTAGT
CTGCTGATAGAAAGCACTATACATCCTATTGTTTCTTTCTTTCCAAAATCAGCCTT
CTGTCTGTAACAAAAATGTACTTTATAGAGATGGAGGAAAAGGTCTAATACTACAT
AGCCTTAAGTGTTTCTGTTCATTGTTCAAGTGTATTTTCTGTAACAGAAACATATTT
GGAATGTTTTTCTTTTCCCCTTATAAATTGTAATTCCTGAAATACTGCTGCTTTAA
AAAGTCCCAGTGTGAGATTATATTATCTAACAATTGAATATTGNAAATATACTTGG
CTTACCTCTCAATAAAAGGGTCTTTTCTATT

~~Sequence ID 717~~SEQ ID NO: 304

TCCACCCACCTTGACCTCCCAAAGTGCTGGGATTATAGGCGTGAGCCACCTCGCCC
AGCCCGATACTAGGACTTATGCAGAAAAACCTTGACATGGAGGAAAGTAAGATCT
AAATAAATACTGTATTCATAGATTAAAAGACTCAGCATAATAAATATAACATTTCT
CCCCAGATTGATGTACAGATTTAACACAATTCCTATCAAGATCCCAGCAAGATTTT
TGTAGATATGTAAAAGATTATTCAAAAATGTAAAAGGAAGGACAAAGGACTAGAAT
AGATAAAACAAAATGGAGAAAGATTTAATAGGAATCACTGTAAGTGAATTTAAGAC
ATACAGAACAATAATAGAACTGCTTGTATTAGTCCATTTTCACGCTGCTGATAAA
GACATACCTGAGATTGGCAATTACAAAGGAAAGANGTTTATTGGCTTACAGTTCCC
ATGGCTGGGGAGGCCT

~~Sequence ID 718~~SEQ ID NO: 305

CTCCTCTGGGTTGAAACCCGGGCGCCGCAAGATGCCGGCTTACCACTCTTCTCTC
ATGGATCCTGATACCAAACTCATCGGAAACATGGCACTGTTGCCTATCAGAAGTCA
ATTCAAAGGACCTGCCCCCAGAGAGACAAAAGATACAGATATTGTGGATGAAGCCA
TCTATTACTTCAAGGCCAATGTCTTCTTCAAAAACCTATGAAATTAAGAATGAAGCT
GATAGGACCTTGATATATACTCTCTACATTTCTGAATGTCTGAAGAACTGCA
AAAGTGCAATTCCAAAAGCCAAGGTGAGAAAGAAATGTATACGCTGGGAATCACTA
ATTTTCCCATTTCCTGGAGAGCCTGGTTTTTCCACTTAACGCAATTTATGCCAAACCT
GCAAACAAACAGGAAGATGAAGTGATGAGAGCCTATTTACAACAGCTAAGGCAAGA
GACTGGACTGAGACTTTGTGAGAAAAGTTTTTCGACCCTCAGAATGATAAACCCAGC

Marked-Up Copy

AAGTGGNGGGCTTGCTTTGTGAAGAGACAGTTCATGAACAANAGTCTTTCAGGACC
TGGACAGTGAAGGGAGCCCGGGCAGCCA

~~Sequence ID 719~~ SEQ ID NO: 306

CGNGGCCGCGTNAACTTTTGATCGTCAGCTGGGGCTGGCAGGCACCTAAATGGGAA
GGGTGATAGCAGTGTGTTGGGGGGAGTTTAGGGAACGGTCCTCTACCGATAGAGGC
AGCANCTCATTGGAATTTCCCTCTGAAGTTGTCTTGCCCCCTGAATCCTGCAGGAA
GGCTGGCAAATGGCCATTTCCCTTCCACTTGAATAGAGACCCATAACTCAAGTATC
TGCCCTTAAGACACCACAGGACTGTTCTTCGCGGGCCCTGCCCCTGGATTTGGGAG
AGGCAGTCCANCTCACCCAACTAGGCTCTGCANGGGGACCANGAGGGATGGGTGT
GTCCACAGGACCAGCCAGACTGATGAGGGATGCGGCAAGCATATTCTCACCACTT
CTTTCACGTTTACAACANACCAGCNTTCCCTGTGTGGCAGGGGTTACATTGGTCAC
CGAGGACCTANAATCATGGAGTGCTCTGGGGATCCGGGCTTGGA

~~Sequence ID 720~~ SEQ ID NO: 307

TCAGTGTTGAATTTTGTGTCAGACACTTTCTCTGCATCAATTGGTATGACCATGTGAT
TTTTTTTCTGTAGCCTGTTAATATGGTTAATTTTCAAATATTGAGCTGATTAATTT
TCAAATATTGAGCTCTCCTTGCACTCTCTGGAATAAGTACCACTTGGTCGTGGTATA
TATTTCTTTTAATATATTGCTGAATTCTGTTTGATCATGTTTTCTTAAAGACTTTC
GTGTCTGTTTTCATGATAGATACTGGTCTATAGTTTTGTTGTAATATCTTGGTTTG
ATTTTGATATCAGGATAATGCTACCTTAATAGAATGAATTGGAGCCAAGTATGGTG
GCAAATGCCTATAGTCCTAGCTACTCAGGAGGCTGAGGTGGTGGGGACTGCTTGAC
CCANGAGTTCAAATCTAGCTTGGGCAATGTAGCAAGAC

~~Sequence ID 721~~ SEQ ID NO: 308

TAGAAGGAATGACTATTCATGTCCAAAGTGAATGGTTTTGTGCAGTGAACAACACA
TGGCGAGGTACTAACTGAGAACTTTTTTCATGCTTTATGCCTACCTCTTGTAGTTG
TTGCAGAGCAAATATAAATTGTAATAAGATAGCTAGGCCTTGCAAGAAACAAACAGA
AAAACCTTAAAAAAAATGATATAAGAGCTGGAGTCTAGTATTTATATGAATCTGTG
AGAGATAATTTTTTTTGGTCTCACTGCAATGAACCAAAGCGGCTGAGTTTGGTTTT
TAATTGTAGCCATGTATTGAAGGCATCTTTTTGACCAACTCTTGTTGGTTCTGTCT
TGAACCATTGTTAATCACTGTGCTGTAATTAGTATAGCTAAATCTTTTCCTTCCTT
GCTCCTCCCCCAGCCCACCCCGTCTTCCCTTAACATTTTTTTCAGGGGGGGTTGGGA
GTGGTTTCATTTTAATGTGAGTGGATGTTTTGATAGTTGTAAGGAAAAAATGCATT
TCAGACACATTTACACATGAGCTATTTTCTTACACAGTATGTCTTATTGGTAATA
AGAATGTAATTCAT

~~Sequence ID 722~~SEQ ID NO: 309

CNTTCCNTAAGAATACAAAAAATTAGCTGGGCGTGGTGGCAGGCGCCTGTAATCCC
ATCTACTCAGGAAGCTGAGGCTGGAGAATCGCTTGAACCCGGGAGGCGGAGGTTGC
AGTGAGCAGAGATCACGCCACTGCAGTCCAGCCTGGGCAACAGTGCGAGACTCTGT
CTCAAAAAAAAAAATAAATAAATTACCTGGGTGTGGCAGCGCGTGCCTGTAATCCCA
GCTACCCAGGAGGCTGAGGCAAGAGAACTGCTTGAACCCAGGAGGCAGAGGTTGCA
TGGAGCTGAGATGGCGCCACTGCACTCCAGTCTGGTGACAGAGTGAG

~~Sequence ID 724~~SEQ ID NO: 310

CTCTCTACTAAAAATACAAAAATTAGCTGGGCACGGNGGTGCATGCCTGTAAACCC
AGCTACCAGGTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCAGGGAGTCGG
AGGTTGCGGCGAGCTGAGATCATGCCACTGCACTGCGGCCTGGAGACAAGAGCAAG
ACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGACNTCACCTAATTGCAG
NGNGNGGACCTTATTTGGCTNTTAATTCAAACCTATTAAAAATGTGAACN

~~Sequence ID 726~~SEQ ID NO: 311

nt:

260

CGGGGTCTGTACCGGGCTGGCCTGTGCCTATCACCTCTTATGCACACCTCCCACCC
CCTGTATTTCCACCCCTGGACTGGTGGCCCCTGCCTTGGGGAAGGTCTCCCCATGT
GCCTGCACCAGGAGACAGACAGAGAAGGCAGCAGGCGGCCTTTGTTGCTCAGCAAG
GGGCTCTGCCCTCCCTCCTTCCTTCTTGCTTCTCATAGCCCCGGTGTGCGGTGCAT
ACACCCCCACCTCCTGCAATAAAATAGTAGCATCGG

~~Sequence ID 727~~SEQ ID NO: 312

CTGAGTNTAGAAATGATGCCATTAATACTGATTGCAAAAACATTACAACCTCAGTAC
TGCAGCTTTCATTCAAATAGGTTATATGTATAAACTGAGTTCAACAATATTGTATT
TGAGATGGTAAAGTTAAAGAAATGCAATAATGTAAATAATACTTAAGAAAATAAGA
TCTCAGGAAACTGTATATACTCTGTACTTTTATGCAACTTTATCAGATCATTTTCAG
TATATGCATCAAGGATATAGTGTATATGACATGAACTTTGAGTGCAAAAACCTGTAC
TATGTACCTTTTGTTTATTTTGCTGTCAACATCTAAATAAAGGTTTTTTTTGTTTGT
TTTTTGTTTTTTAATTGTTTTGTTTTAAAGATTGTTTTAATTAATTAATAAATA
ATTGTTTTTAATTAACAATTGTTTAATTGTTTTAAAGTCGCCAGGCTGAGGCAGGT
GAATCACAAGCTTAGGAGTTGGAGGCTAGCCTGCCAACATGGTGAAACCCCGTCTC
TACTAAAAATACAAAAAATTAACCTGGGTGTGGG

Marked-Up Copy

~~Sequence ID 728~~ SEQ ID NO: 313

CCCATCTGCACCAGTACACAGGCAGGCATTATCATTCTTCACCTACTTTTTTAAATA
GTGGCAACTTGGGATTCATTCTGGTGATTCTGAACCTTGCCTCATAGCTTAAAGTA
TAAAAAAGATTCAAGAGCAGTGAGGTTTGTTCCTTTCCAGTGAATGGTGGACTGAGT
GGTGCGAGGTGGAGGGCTAACAAGAGGAAAGAACTACATTCTTCAGAATACAGTGA
TGAAAATTCATTTTGAACCTCAAATATTTTCATTTTGGATATTCTCCTGTTTTTAT
TAAACCAGTGATTACACCTGGCCATCCCTCTAAATGTTCTAGGAAGGCATGTCTAT
TGTGATTTTGTATGAAGACAGAATTATTTTTCTCTGTAGAAACACAGATACCACTTT
ATCAGGGGAAGTTAGTCAAATGAAATGGAAATTGGTAAATGGACAAAAGCTAGCTA
GTAAAAAGGACGACCCAGCAACATGCTTTAACCCCATTTGTATGTTTGTGGAAAGAG
CATAGTTTAACATCTTGAGAAATTTGGGACATAAAAGTTTTTCATNGGTAGACAGTT
CATGGCAGTATATGAATTGACATAATGGAAATAATCTGATTTTATTTTTTACAATA
ACATCCTTTCCCC

~~Sequence ID 736~~ SEQ ID NO: 314

nt:

641

GGAATTCCAAGTGCTTGGGGATAATGATACCTCTGACCTTTCTTCCTTTTGGGAAG
TACTTGAGTGTGCAGCTGCATGAGGCCTCAGCAGGAGAGAGATTTTAGGTCCAAGA
AGCTATACCAGTAGGACAAGGCAGGAAAATACTACACTTTCAGGATCAAGCCCCCTC
TGACTCTCATTGTGAAACTGGATGTTTGCTAAGCACCTGCTTCTTAAGGATGCCGA
GGGATTTAATGATACTCCCAGAAACCTGGAGAGATTAATGGGGCCTATGGAGAAGT
GCTCTGAACTCAGTGTTGGGACTTGAATAAAATTAACCATTTGTCATGTTTTCAGAA
CAACTAAGCTGTTTTATATTTTCATGTGCATGAAAGCCCTAGAATAAGTTGTGTTA
TTTCCAGAAATGAAATAGATCCCACAGTTAGATGATGTGGCCATTAGGAAGTACCA
AATTTATAAAAAATCACTGGAGGTCTGTCTGAGCAGTACCTAATAAAATATAGTATA
CTGAAAGTGAACAGATACTTTGTCTCTTTCTTTGGCTGCTTGATCTTTATCTGTGT
CTGCCGTACAGTGCACCCTTAAAGTATTCTACACCAGTGCTTCTCAAACCTGGAAAT
GTGCATGTAAGTCACCCANGGGTCT

~~Sequence ID 739~~ SEQ ID NO: 315

TGCATGCCCATAGTCCCAGCTATTTGGGAGGCTGAGGCAGGAAAATCGCTTGAACC
CGGGAGCCAGAGGTTGCAGTGAGCCGAGATCGCACTCCAGCTTGGCGACAGAACAA
GACTCTGTCTCAAAAAAAAAAAAAAAAAAGAAATCTTGGGATCCTGAACCCCTTACTC
GAAGGGCTAAGGTAGCATCTCAGCATGTCTTATTCGAGACTTCGTANAACCAGACC
TGCTGTTTGTAGATGTTAATTAATCAAACCTTTCTCTACTCATTCTGGACCAGTTA
AGGTTTTCTCCTTCTCCGTATGAGTTTTGATTTTCGTCCTCCTTGGTTGGAGATCA

Marked-Up Copy

CACTTTGGTCTGCTGCTAAGTTGGATGCCTCCCCTGTCTTTCCCTAAGTCTAGGG
CTTCANACCCCAGTGTGGGGAGAGGGACTTTCGTTTCCTGCCCCCTCACCACATCAG
ACACAGGCAGGCAAGAATAAGATGGCCAAAAGGCCGATGAACTTCTTGACCTAGCC
TGGGACATTACCTGTTACTAGGTGGACTTCACTGCCTGTGAATGGAAGCTGAAGGG
CTGTTTTTTTTGGTTTGTATTTGGACAGGCCAGGCTTANAGAGGGAGAGAACTGGGC
TACTCTTCAGCAGTGATCTTTAAATGCC

~~Sequence ID 747~~SEQ ID NO: 316

CAGAGTGCAAGACGATGACTTGCAAAATGTCGCAGCTGGAACGCAACATAGAGACC
ATCATCAACACCTTCCACCAATACTCTGTGAAGCTGGGGCACCCAGACACCCTGAA
CCAGGGGGAATTCAAAGAGCTGGTGCGAAAAGATCTGCAAAATTTTCTCAAGAAGG
AGAATAAGAATGAAAAGGTCATAGAACACATCATGGAGGACCTGGACACAAATGCA
GACAAGCAGCTGAGCTTCGAGGAGTTCATCATGCTGATGGCGAGGCTAACCTGGGC
CTCCCACGAGAAGATGCACGAGGGTGACGAGGGCCCTGGCCACCACCATAAGCCAG
GCCTCGGGGAGGGCACCCCCTAAGACCACAGTGGCCAAGATCACAGTGGCCACGGC
CACGGCCACAGTCATGGTGGCCACGGCCACAGCCACTAATCAGGAGGCCAGGCCAC
CCTGCCTCTACCCAACCAGGGCCCCCGGGCCTGTTATGTCAAACCTGTCTTGGCTGT
GGGGCTAGGGGCTGGGGCCAAATAAAGTCTCTTTCCTC

~~Sequence ID 757~~SEQ ID NO: 317

nt:

583

GAACCCTGCGGAGGGACTTCAATCACATCAATGTAGAACTCAGCCTTCTTGAAAG
AAAAAAAAGAGGCTCCGGGTGACAAATGGTGGGGTAACAGAAAGGAAGTGGCTAC
CGTTCGGACTATTTGTAGTCATGTACAGAACATGATCAAGGGTGTACACTGGGCT
TCCGTTACAAGATGAGGTCTGTGTATGCTCACTTCCCCATCAACGTTGTTATCCAG
GAGAATGGGTCTCTTGTTGAAATCCGAAATTTCTTGGGTGAAAAATACATCCGCAG
GGTTCGGATGAGACCAGGTGTTGCTTGTTTCAGTATCTCAAGCCCAGAAAGATGAAT
TAATCCTTGAAGGAAATGACATTGAGCTTGTTTCAAATTCAGCGGCTTTGATTCAG
CAAGCCACAACAGTTAAAAACAAGGATATCAGGAAATTTTGGATGGTATCTATGT
CTCTGAAAAAGGAACTGTTTCAGCAGGCTGATGAATAAGATCTAAGAGTTACCTGGC
TACAGAAAGAAGATGCCAGATGACACTTAAGACCTACTTGTGATATTTAAATGATG
CAATAAAAGACCTATTGATTTGG

~~Sequence ID 758~~SEQ ID NO: 318

nt:

424

Marked-Up Copy

CTTGGCTCCTGTGGAGGCCTGCTGGGAACGGGACTTCTAAAAGGAACTATGTCTGG
AAGGCTGTGGTCCAAGGCCATTTTTGCTGGCTATAAGCGGGGTCTCCGGAACCAAA
GGGAGCACACAGCTCTTCTTAAAATTGAAGGTGTTTACGCCCAGATGAAACAGAA
TTCTATTTGGGCAAGAGATGCGCTTATGTATATAAAGCAAAGAACAACACAGTCAC
TCCTGGCGGCAAACCAAACAAAACCAGAGTCATCTGGGGAAAAGTAACTCGGGCCC
ATGGAAACAGTGGCATGGTTCGTGCCAAATTCGAAGCAATCTTCCTGCTAAGGCC
ATTGGACACAGAATCCGAGTGATGCTGTACCCCTCAAGGATTTAACTAACGAAAA
ATCAATAAATAAATGTGGATTTGTGCTCTTGT

~~Sequence ID - 764~~ SEQ ID NO: 319

nt:

626

GATTTTTTTTTTTTTTTTTGAGATGGAGTCTTCTCTGTGCGCCAGGCTGGAGTGCA
GTGGTGAAATCTCGACTCACTGCAACCTCCGTCTCCTGGGTTCAAGCAATTCTCCT
GCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCACCAGCCACCACGCCCCGGCTAAT
TTTTGTATTTTAGTAGAGACAGGTTTTACCATGTTGGCTAGGCTGATTTTGAAC
TCATGACCCCAAGTGATCTGCCCCGCTCGGCCTCCCAAAGTGCTGGAATTACAGGT
GTGAGCTACCACTCCCAGCCAATGATTACATTTATAAGGTAAAATAACTTGTGCCA
ATCTGTACAAGTGAATTCAGATTTAAAATTTTAATTGTAAAAAGATATCCAGGTGA
TATTTCTCCCTGAATAATTTAGTTTCCTTTTCTATTTCTTGATATAAAAGTACTCA
GCATTGAAGTAATTGCTATCTTCACATTTCTTCCTATTTGAGCTGTCTAAATAAGT
AGTCCTACATATTTTCCCCCAACACAAAAAACCCAGAAAAGAATTATTTTATACT
GGATTTTTTTGGTTGTAGCAGGAACCTAAAGNGCCAATTGTAACATGCATGTTCT
TTTTGGCAAA

~~Sequence ID - 766~~ SEQ ID NO: 320

GTCCATCCTGCAGGCCACAAGCTCTGGATGAGGAACTTGAGGCAAGTCACCAGCCC
CTGATCATTTTCGCCTAAAAGAGCAAGGACTAGAGTTCCTGACCTCCAGGCCAGTCC
CTGATCCCTGACCTAATGTTATCGCGGAATGATGATATATGTATCTACGGGGGCCT
GGGGCTGGGCGGGCTCCTGCTTCTGGCAGTGGTCCTTCTGTCCGCCTGCCTGTGTT
GGCTGCATCGAAGAGTAAAGAGGCTGGAGAGGAGCTGGGCCCAGGGCTCCTCAGAG
CAGGAACTCCACTATGCATCTCTGCAGAGGCTGCCAGTGCCAGCAGTGAGGGACC
TGACCTCAGGGGCAGAGACAAGAGAGGCACCAAGGAGGATCCAAGAGCTGACTATG
CCTGCATTGCTGAGAACAAACCCACCTGAGCACCCCAGACACCTTCCTCAACCCAG
GCGGGTGGACAGGGTCCCCCTGTGGTCCAGCCAGTAAAAACCATGGTCCCCCCT
TCTGTGTCTCAGTCCTCTCAGTCATCTCGAGCCTCCGTTCAAATGATCATCATCA
AACTTATGTGGCTTTTTGACCTTTGAATAGGGAATTTTTTAAAATTTTTTAAAAA

TT

~~Sequence ID 768~~ SEQ ID NO: 321

CCAGCGCAGGGGCTTCTGCTGAGGGGGCAGGCGGAGCTTGAGGAAACCGCAGATAA
GTTTTTTTCTCTTTGAAAGATAGAGATTAATACAACACTTAAAAAATATAGTCAA
TAGGTTACTAAGATATTGCTTAGCGTTAAGTTTTTAACGTAATTTTAATAGCTTAA
GATTTTAAGAGAAAATATGAAGACTTAGAAGAGTAGCATGAGGAAGGAAAAGATAA
AAGGTTTCTAAAACATGACGGAGGTTGAGATGAAGCTTCTTCATGGAGTAAAAAAT
GTATTTAAAAGAAAATTGAGAGAAAGGACTACAGAGCCCCGAATTAATACCAATAG
AAGGGCAATGCTTTTAGATTAAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTT
TAAAAGTTGTAGGTGATTAAAATAATTTGAAGGCGATCTTTTAAAAGAGATTAAA
CCGAAGGTGATTAAAAGACCTTGAAATCCATGACGCAGGGAGAATTGCGTCATTTA
AAGCCTAGTTAACGCATTTACTAAACGCAGACGAAAATGGAAAGATTAATTGGGAG
TGGTAGGATGAAACAATTTGGAGAAGATAGAAGTTT

~~Sequence ID 773~~ SEQ ID NO: 322

GAGGAAAGGGGAGTTAATATTTAGTGGACAGAATTCAGTTTTACAGATGAAAAGA
GTTCTGGAGATAGACGGTGTTGATAGTTGCACAGCAGTGTGAATGTGCTCATTGTT
ACCGAACTTAAAAATGTTTAACATAGTATTATGTGATTTTTTATTTTGCCACTTAAA
AAAAAAGAATGAAGTACTGATACATGCTACAACATGGGTGAGCTTTAAATACATTC
TGCTCAGTGAAATAAGCCAGATGCAAAAGATCACATATTATATAATCCACTTATAC
GAGATACCTAGAATAGGCAAATTCATAGAGACAGAAAGTAGAATAGTGGTTCCCAG
GGGCTGGGGACAAGGGGGCAGTGAGAGATTGAGAGTTATTATTAATGCGTACAGAG
TTTCAGTTTGGGCTGATAAAAAAGTTCTGAAGATGGATGGTGATGATGGTTGTACA
TCAATGTGAGTGTAATTACCGCCACTGAACTGCCCTTAAAACGTTTAAAAGAGTA
AATTTTATGTTGNGTATATTTTACCATAAT

~~Sequence ID 776~~ SEQ ID NO: 323

TTTTTTTTTTCATAAGAGGCAAGTACAAGAAAAAGCTTAATTACTTTAACTTCTAAG
TAGTTTGGAACTCTAAATAAATAGGAGTTACCAAATATATGCGCTTCTGTGAATAGT
TTTCCCCCACATGTTTATTTATATTTTTTGCATCTCATCAAACCTAACAGATTCTAA
AGTCTCTGGTGATAATGACAATATCTGCTACGGAGAGACTAGCCTGGGGGAAGAGG
ATCTCCCTGAACAAGGATAGCGGAGTTGCTGCAGCTTTCAAATGAAGCTGGACATT
TAGCTGCGGGGGTAGCACCCCTTTGATCAAGGCAGCCCAAAGATGAGTTTCAGGGAT
GGGACTGACAGAAGAGAAAAGTTCTTCCCAGCCCTTTCTACTTTTTTCTCTTTGTTT
CTCAGGCTTCTGGCCGTCTTCAGTTTTTACAAGTTTCACTCTCAACCCTAAACAGT

Marked-Up Copy

ACTTCTGTGAAGTACCCTTTGGCCCCTCGTTTTTCAGCTCCTAAACTCACCTGGAAA
TAGATGTCAATCTAATTTTGGGTCTGACTAGTGCAGTAGGCATTTTTTGGTGA

~~Sequence ID 782~~SEQ ID NO: 324

CTCACACAGAACAAAAATGAATGAGTGTGGCTGTGTGCCACTATCACTGTGTCTAC
AAAAACAGCCAGTGGGCCTGATTTGGCCCTTGGCTGCAGTGCGCCCGTCTCTGTTT
TTGAGGAATAAAATCGCATCATTTTCATATGGCTAATGCAATTTTTTTTCCCATCTGG
AAGCAACATCTGATTGGACTCATCTTGTATGGTGCTTGTTACAGTCTCTGTAAATG
GGAGAGGGTCCGAGAATAGCTCTTCCTGTTTTTCATCAGGACTGTTTTTAGGGATGG
CAAAGAAGTCAGTGTGTCCAGCCTGTGTCCTCCTCACCACGTGGCTGATTCCTGAA
TCTGCATGTGCANCACNTGCCGTTGTCTGGGGCATGATCTGTGTGA

~~Sequence ID 785~~SEQ ID NO: 325

nt:

556

CTTTTCTCTGGGTATAGATTTACCCTAGCACCTATCTCATTATATTGAATTTTCCA
GCATATTTTAAATAAACTATTAATTAGTCACACTATTTCTTAAAAGTCACACTATCA
ACTAATCGTGACCGCAATTATCTAGGGGTGATAATCTGCTGAGTCTACTCTTTAAA
TACACTGGGACCCAGCATATTGAGTTATATTGGCACAGAACTTCACTCTGGGTAT
AGATTTACCCTAGTACCTTGCCGGCAGGATCCTATTATTCATGGTTGTACAAGCAA
GGTTCAGGGAAGAGGCTGGCACAGAGAAGGTACCTGGTAACTGTTGTTTGAGGCTG
AATTCAGCTCAACTCAGCTCCAGTAGAGATGGTGTCCCCTTCTCTACCGTGTTGAG
ATAGTGTGCAGTCCCTTCCTAAGGGCTGTTACCCACCGCAATAGGACTTGTGAGCT
TCAACTTTTAAATTTCTCTGCTCCCGCTGGGACCCACCCGCTTCAAAAATCATCAT
GGNGGNTTTAGCACCAATTTAGTAAACACAACTGTCTGAAATATTTTGGAT

~~Sequence ID 796~~SEQ ID NO: 326

GAACATTCAAGATAGTGAGAGGAAGAAAAAGATATGGCTGTACGGGACCGAGGTCT
CTTCTATTATCGCCTCCTCTTAGTTGGCATTGATGAAGTTAAGCGGATTCTGTGTA
GCCCTAAATCTGACCCTACTCTTGACTTTTGGAGGATCCGGCAGAAAGACCTGTG
AATAGCTGGGCCTCAGACTTCAACACACTGGTGCCAGTGTATGGCAAAGCCCACTG
GGCAACTATCTCTAAATGCCAGGGGGCAGAGCGTTGTGACCCAGAGCTTCCTAAAA
CTTCATCCTTTGCCGCATCAGGACCCTTGATTCCTGAAGAGAACAAGGAGAGGGTA
CAAGAACTCCCTGATTCTGGAGCCCTCATGCTAGTCCCCAATCGCCAGCTTACTGC
TGATTATTTTGAGAAAACCTTGCTTAGCCTTAAAGTTGCTCATCAGCAAGTGTTGC
CTTGGCGGGGAGAATTCCATCCTGACACCCTCCAGATGGCTCTTCAAGTAGTGAAC
ATCCAGACCATCGCAATGAGTAGGGCTGGGTCTCGGCCATGGAAAGCATACCTCAG

Marked-Up Copy

TGCTCANGATGATACTGGCTGTCTGTTCTTAACAGAACTGCTATTGGAGCCTGGAA
ACTCAGAATGCAGATCTTTTGTGAACAAAATGAAGCAAGAACCGGAGACNCTGAAT
AGTTTTATTTCTGTATTAAAACTGNGATTGGAACAATTGAAGA

~~Sequence ID 801~~ SEQ ID NO: 327

CCACTCCACCTTACTACCAGACAACCTTAGCCAAACCATTACCCAAATAAAGTAT
AGGCGATAGAAATTGAAACCTGGCGCAATAGATATAGTACCGCAAGGGAAAGATGA
AAAATTATAACCAAGCATAATATAGCAAGGACTAACCCCTATACCTTCTGCATAAT
GAATTAAGTAGAAATGAGGATTCTGACCTTGACTTTGATATCAGCAAATTGGAACA
GCAGAGCAAGGTGCAAAACACAGGACATGGAAAACCAAGAGAAAAGTCCATAATAG
ACGAGAAATTCTTCCAACCTCTCTGAAATGGAGGCTTATTTAGAAAACAGAGAAAAA
GAAGAGGAACGAAAAGATGATAATGATGATGAGTCAGGTAAAAGTTCCAGAAATGT
GAACAACAAAGATTTTTTTTGATCCAGTTGAAAGTGATGAAGACATAGCAAGTGATC
ATGATGATGAGCTGGGTTCAAACAAGATGATGAAATTGCTGAAGAAGAAGCAGAAG
AAGGAAGCATTCTGAAATATGAATGAAAAAAATTACATCTTTAGAAAAAGAGTTA
TTAGAAAAAAGCCTTGGCAGCCGTCNGGGGGAAGTGACGCACAGAAGAGACCAGAG
AATAGCTTCCTGGANGAGACCCTGCACTTTACCCATGCTGCTGGATGG

~~Sequence ID 808~~ SEQ ID NO: 328

nt:

641

CCGGGTTTTAGTATTTAAACCAAGAGCCTTTTAAATATTGAAAACCCATAGTTCAGA
AAATGTTAGTATTGCTGCCCTTCTTCACATAAAATTTTTTTTTTAAATTATACTATTA
TTTTGCTTAATTTTATATTGGGTAAAACAACCTTCAAGAAGGTTAACTAGGAAAG
AAGACCTTTTTGTTTTATTTTTACTATTTATATATAGAAGACAAATCAGCATTG
TGATAGTTTTACATGACCAGTTATCAAACGGTCATAGTATGAAGTGTGCAGTTGTT
CATTATTAGTAAATTATGTTTGATTTTTTAACTATTTAGTACTAATAGTTGAGATG
AAAAGTGAAGAAAAATGCCAATGTGACGTTTGTGTATAGCTAGCCTTAAAAAAGT
CCCATGTTTTTTAGGTGACTTTTTTCCCCCTCTTAGTACTCTGGAGAAACAATGAAG
ATGGGCCATCTCAATTCCAGATGTAAACAAAAAGTAATTTTTATTTCAACATTTAA
TGTAAGTCTATTATTGNGGATTCTTGNCTTGNGTATTTTCTTTCCCTTATTCAG
TAATATAGAATAACTTTCCTTAAATGATTTGATCCAAGATACGTCATTTCTGTAT
TGGCAAAATGCCNCTATTAAAGTGT

~~Sequence ID 814~~ SEQ ID NO: 329

nt:

132

GTAAAGTGATACATTTTTTATACCAAATGTGTTTATTTTTTTGTGCAAGTAATCCT

Marked-Up Copy

TAAAATTGCAATTGTATTAGGTGTTAAAATAAAGTTTTTAAAAAATTAAAAAAA
AAAAAAAAAAAAAAAAAAAA

~~Sequence ID 817~~SEQ ID NO: 330

GACAACCTTAGCCAAACCATTTACCCAAATAAAGTATAGGCGATAGAAATTGAAAC
CTGGCGCAATAGATATAGTACCGTAAGGGAAAGATGAAAAATTATAACCAAGCATA
ATATAGCAAGGACTAACCCCTATACCTTCTGCATAATGAATTAAGTAACTAGAAATAACT
TTGCAAGGAGAGCCAAAGCTAAGACCCCGAAACCAGACGAGCTACCTAAGAACAG
CTAAAAGAGCACACCCGTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGG
CGACAAACCTACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCA
ACTTTAAATTTGCCCACAGAACCCTCTAAATCCCCTTGTAATTTAACTGTTAGTC
CAAAGAGGAACAGCTCTTTGGACACTAGGAAAAACCTTGTTAGAGAGAGTAAAAA
TTTAACACCCATAGTAGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTC
AACACCCACTACCTAAAAAATCCCAAACATATAACTGAACTCCTCACACCCAATT
GGACCAATCTATCACCTATAGAAGACTAATGTTAGTATAAGTAACATGAAAACAT
TCTTCTNCGCATAAGCCTGCGTCAGATTAAACACTGAACTGACAATTAA

~~Sequence ID 821~~SEQ ID NO: 331

nt:

370

AAAGAGCTCCCAAATGCTATATCTATTTCAGGGGCTCTCAAGAACAATGGAATATCA
TCCTGATTTANAAAATTTGGATGAAGATGGATATACTCAATTACACTTCGACTCTC
AAAGCAATACCAGGATAGCTGTTGTTTCANAGAAAGGATCGTGTGCTGCATCTCCT
CCTTGCGCCTCATTGCTGTAATTTTGGGAATCCTATGCTTGTTAATACTGGTGAT
AGCTGTGGTCCTGGGTACCATGGCTGGTTTCAAAGCTGTGGAATTCAAAGGATAAA
TTAATGAAGAAAACAAGCGGAGCTGAAGAAGAAAGTACAATATGGTGCTGTCTTCC
TAATGAAATAAATTCACTAAATGGACATTAAAA

~~Sequence ID 825~~SEQ ID NO: 332

AGACTCGAGCAAGCTTATGCATGCATGCGGCCGCAATTCGAGCTCGGCCACTTGGC
CAATTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTC
GTGACTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCGAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTT
GCGCAGCCTGAATGGCGAATGGAAATTGTAAGCGTTAATATTTTGTAAATTCG
GTTAAATTTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAA
TCCCTTATAAATCAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGG
AACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGT

Marked-Up Copy

CTATCAGGGCGATGGCCCACTACGTGAACCATCACCCCTAATCAAGTTTTTTGGGGT
CGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCGATTTAAAGCT
TGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAAAAGCCAAANGGAG
CCGGCGCTAGGGCCTGGCAAGTGTACGGGCACGCTGCGCGTAACCACCCACACCCC
GCCGNGCTTAATGCCCCNTTCAGGGCGCGTNCTGATGCCGNATTTTNTCTTACNCA
TNTGTGCNGGNTT

~~Sequence ID 833~~SEQ ID NO: 333

TAAAATAATGGCAAAAAACAAACAAAAACAAGTTCTCTAAACAGAAAGGAAATTA
CTAAAGAAGGAATCTTGAAATAACAGGAAAGAGGAAATACCACAGTAGGCAACATT
ATGGGTAAATAAAACAGACTTTCCTTCTTTAGTTTCCTAAAATATGTTTGATGATT
AATGCAAAAATTACAATATTTTCTTATGTAGCACTAAAGGTATGTAGAGAAAATAT
TTAAGATAATTGTACTGTAAGCGGGAGATGACAGTGACATAAAGGCAACGTTTTTA
TACTTCACTCAAACCTTTATGTATTAATGTAATCCATAAAGCAACCAAAAAAGCTAT
ACTAAGTACATTCAAAAACACAATAGATAAACCAACAAAATTCTAAAGGATGTAC
AAGTAACCCACTGGAAGCTGCAAAAAATGTAAACAGAACTAAAAACAGAGAATAA
ATGAAAAATTAAAAACGAAATGGCAGACTTAGGCCCTAATATACAAATTATCACAT
TAAATATAAATGGTCTAAATACCAACTGTAAGACAGAGATTAGCAAAGTCGATT
TAAAAACATGACTCAACTACGTGCTGTCTACAAGAACTCACTTCAAATATACCAA
GATAGGAAGGTTGAAAGTAAACGATGGAAAAAGATGTATCATGTGAACATTAATC
AAAGGAAAGCAGGGGTGGCTATATTAACATCAGGTAAAATAAACTTT

~~Sequence ID 837~~SEQ ID NO: 334

nt:

603

TGAGGNTGGTCATGATGCANAAGCTACTCAAATGCAGTCGGCTTGTCTTGGCTCTT
GCCCTCATCCTGGTTCTGGAATCCTCAGTTCAAGGTTATCCTACGCGGAGAGCCAG
GTACCAATGGGTGCGCTGCAATCCAGACAGTAATTCTGCAAACCTGCCTTGAAGAAA
AAGGACCAATGTTCTGAACCTACTTCCAGGTGAATCCAACAAGATCCCCCGTCTGAGG
ACTGACCTTTTTTCAAAGACGAGAATCCAGGACTTGAATCGTATCTTCCCACTTTC
TGAGGACTACTCTGGATCAGGCTTCGGCTCCGGCTCCGGCTCTGGATCAGGATCTG
GGAGTGGCTTCCTAACGGAAATGGAACAGGATTACCAACTAGTAGACGAAAGTGAT
GCTTTCCATGACAACCTTAGGTCTCTTGACAGGAATCTGCCCTCAGACAGCCAGGA
CTTGGGTCAACATGGATTAGAAGAGGATTTTATGTTATAAAAGAGGATTTTCCCAC
CTTGACACCAGGCAATGTAGTTAGCATATTTTATGTACCATGGNTATATGATTAAT
CTTGGGACAAAGAATTTTATAGAAATTTTAAACATCTGAAAA

~~Sequence ID - 839~~ SEQ ID NO: 335

nt:

71

ATTTATCTAATATTTGGTTTAATAAAATGTGAATAATGAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

~~Sequence 849~~ SEQ ID NO: 336

nt: 622

TTTTTTTTTATTTTTTGAGAATGGAGTCTTGCTCTGCCGTCCAGGCTAGAGTTCAG
TGGTGCATCTCAGCTCACTGCCACCTCACCTCCTAGGTTCCAGAGATTCTTGTGC
TTCAGCCTCCTCAGTAGTTGAGAATACAGGAACACGCCACCACGCCTAGCTAATTT
TTGTATTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAATC
CTGGCCTAAGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTATAGGCGTG
AGTCATTGTCCCCAGCCGGATGTTTTTCATCTTGATTTGCCTTAGTTTCTAAATCTC
ATCCTCTCCATTTTCTCCTGTTAGTAGTCACAGAGAACCAAATTCGTCAAGTTAT
GAACTAAAGTCTCTCTTCCACAAGTCTTCCTGTGTTCTGCCTCAAGTGAAGTTGA
AAGAACATCAGTTTGTGGGAAGGTTGAAGACCGAATGATCTGCTGGGAAATCACTG
AGGCATTGCCATTCTCTTGAGGAATTTTCATTTTCATCGAAGTTTCGGTTTATATCC
CTTTCTTGGTGAGTACTATTGCTGTTATGTAAATTAAATGAGTCGTCATCCTTCTT
NTGAGC

~~Sequence ID - 860~~ SEQ ID NO: 337

nt:

501

GTGAAATCACTTTCATGGATTATTAATGGATTTAAGAGGGCATCAATCAGCTCAAC
TCAAGATTTTATAATCATTTTTTAGTATTTAGATTGTGCCTCAAAGTTGTAGTACCT
CACAATACCTCCACTGGTTTCCTGTTGTAAAAACCTTCAGTGAGTTTGACCATTGT
GCTCTTGGCTCTTGGGCTGGAGTACCGTGGTGAGGGAGTAAACACTAGAAGTCTTT
AGTACAAAAGTCTCTAGGGACACCTGGTGATTCCTACACAAGTGATGTTTATATT
TCTCATAAAGAGTCTTCCCTATCCCAAGGCTTCATGATGCCAGTAGCCATATATG
ATAAATTATGTTTCAAGTGATAACTTAGTTATCAGAAATCAGCTCAGTGGTCTTCCCC
GCCATGATTCACATTTGATGAGTTTTTAAAAATCAAAGTGATTTTGAAAATCTCTA
ATGGCTCAGAAAATAAAAACATCCAGTTTGTGGATGACTATATTTAGATTTCT

~~Sequence ID 864~~ SEQ ID NO: 338

TTGTGTTTTTAGGACTCCTTATCTAAATTAAGGCAGAGAAGTTACAGTATTTATAT
CTGCATTAAATCTCAATTCCAGAAAAACCTTTTGAAAAATTATTTAATCCTCTGGA
AACTATTGATATGATACAGGAGAAATTTTCAGAAGTTTATTGAATAATTTAATATC
ATTTAATAGGACACTCTGGCTTGTATATAAGCAGATACGTTACTCAGACTTCTTGG

Marked-Up Copy

CTGTACTCTAAAATAATATATGTACTAGTCTCCTAAATATTACTAGCTCACCTTTC
AAAATGCATACTAATATTTCAATGTCTTTCTTCAATTTGAAAAGCTCTTGAATATC
TACTTGTGATAGCCCTAAGAGCTGAGATAATTATTTCCAGGAGGTTGAATCCCTGA
TTCTTAACTGTTTCAAGCAATGCATAAGCAAGAGAGAATATGACATAAGAGGACCATT
TCTACATTAGCCATTTTTTTTTTCAACAAGATACCTATGTGAATACAGGGCACCTGGGA
GGGTAAGTGGAGGACTATTTCTAACTATATTTATAAGCACATACTGATATTGGTGA
ATCAAAACCTACAGCAGTGCTTCTCAGATGGGAAGGGAGACAATGTGTAAGGAGAT
CAGGAATTCATTAG

~~Sequence ID - 865~~ SEQ ID NO: 339

nt:

122

CCANAATCCACTCTCCAGTCTCCCTCCCCTGACTCCCTCTGCTGTCCTCCCCTCTC
ACGAGAATAAAGTGTCAAGCAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

~~Sequence ID - 867~~ SEQ ID NO: 340

TTTTTTTTTTTTTTTTTTTTTTCAGAGTCACAGATATTGTATAGCTGAGGTAAGCATTT
TACAACCTTTTCAGACACAAGTAAGTACATAAATATTATTTTACAACCAACAATNTT
TAATATTTCCACATTGAANAATAGATGTGATAATTAAATCTTTTATAAGGTTTTAA
AAAGACATGAAACATAAACCTAATTATACATAAAAGAAAAGAATTTTAAACAAGAG
CTTATTGNGATGACATTACTCATAACTTTTACCTTTAAAACCTTTTCTTGGGTAGC
TATTCAAAAGTAAAGACCACAAGTTTTGTTGCCCANATTTCTTATGTTTNGTATAT
TTAAGCTCTTTATTTATTGAACAGATGNGTCATTAATTCATTNGGAGCATTACTAT
TATCAGTAAAATTTGATTTTTTTTTTCCCCTCAGTCATAGGTAAATCAGCTCCACCT
GGAATTTCTAAGGACCCAGTTTTAGTCAATATTTTCAAGTAATCATGACCTCAGAA
ATAGTCTTAATTAAGATAACAAATATTAGCCATCAAAATGGAACCAAGACAAGATT
CTAATGTTTGTAACAGTCAATCCATATTTATGAATATTAGCATATATTGGNGAAT
AGTTAAGGCAAAAGGGTCTAGCAG

~~Sequence ID - 869~~ SEQ ID NO: 341

nt:

667

TTGTGTTTTTAGGACTCCTTATCTAAATTAAGGCAGAGAAGTTACAGTATTTATAT
CTGCATTAAATCTCAATTCCAGAAAAACCTTTTGAAAAATTATTTAATCCTCTGGA
AACTATTGATATGATACAGGAGAAATTTTCAGAAGTTTATTGAATAATTTAATATC
ATTTAATAGGACACTCTGGCTTGTATATAAGCAGATACGTTACTCAGACTTCTTGG
CTGTACTCTAAAATAATATATGTACTAGTCTCCTAAATATTACTAGCTCACCTTTC

Marked-Up Copy

AAAATGCATACTAATATTTCAATGTCTTTCTTCAATTTGAAAAGCTCTTGAATATC
TACTTGTGATAGCCCTAAGAGCTGAGATAATTATTTCCAGGAGGTTGAATCCCTGA
TTCTTAACTGTTTCAATGCATAAGCAAGAGAGAATATGACATAAGAGGACCAT
TCTACATTAGCCATTTTTTTTTCACAAGATACCTATGTGAATACAGGGCACCTGGGA
NGGTAAGTGGAGGACTATTTCTAACTATATTTATAAGCACATACTGATATTGNTGA
ATCAAAACCTACAGCAGTGCTTCTCAGATGGGAAGGGAGACAATGTGTAAGGAGAT
CAGGAATTCATTAGTCACCTTTCAGATGGTTTAATGCATACAGCTGTACCG

~~Sequence ID 870~~SEQ ID NO: 342

GGAGTTTGAGCAGATCCTTCAGGAGCGGAATGAACTCAAAGCCAAAGTGTTCTGC
TCAAGGAGGAACTGGCCTACTTCCAGCGGGAGCTGCTCACAGACCACCGGGTCCCC
GGCCTTCTGCTCGAGGCCATGAAGGTGGCTGTCCGGAAGCAGCGGAAGAAGATCAA
GGCCAAGATGTTAGGGACACCAGAGGAAGCAGAGAGCAGTGAGGATGAGGCTGGCC
CATGGATCCTGCTCTCCGATGACAAGGGAGACCATCCCCACCCCCGGAGTCCAAA
ATACAGAGTTTTCTTTGGCCTATGGTATCGGGGTAAAGCTGAATCCTCTGAGGATGA
GACCAGCAGCCCTGCACCCAGCAAGCTAGGGGGGAGAAGAGGAGGCCCAACCACAGT
CTCCAGCTCCTGATCCGCCCTGTTCTGCCCTCCACGAACACCTTTGTCTGGGGGCC
TCAGCCGCCCCAGAGGCCTGACTTAGGGGTCTGGCTGTGGAAGGATGTGTGGCCTC
AAATGAGGACAGGGCTCCCGCCTTCACAGCCCTCGCCAGGGGTCTGCCCCAATCCT
GGCCTGCATCAGGCAAGGACGGGGTCTCAGC

~~Sequence ID 871~~SEQ ID NO: 343

nt:

642

GCAAGTCTTCAGTATGTACATTTATCCCCTAGAAGAAGAAAAATTAGTTGTGCATG
AAAAAGAAACATTAACTGCAAAGCTAAATGCTCACACTCTAAATCAGTGCTCTCCA
AAGTACAGCAGGCGGGAAAAGAAAATGGTAGATTTTTTTTCTTCCAATTACTTTAAC
TTATTCTTTTTAATGGACACTTCATACATAAATATATTCACAATATATTAATATAT
ACATAATGTATAAGCATAACATATTGAATGTGCAGTCAAAAAATGTACTAATGGAAT
GCTCTACCAAAACAAGTTCACGTTTCTGTAAAATGGGAATAATTTTTTAAAAG
GCATACAGTCTGAACATTTTTTAGATTATTCATAAAATCTATTCAGAAAGTTAACT
AAAAAATTTAACGTATGCCTATAACAAATTTTGTACTTAATGTAATTGNTTTTCAT
CCTGAGATCTAATATCCTCGTTTTTAAGTAGAGCCACTTGTTTGCTACAGTTTAGT
CAAAACGTTAACATTAGATGGGTAAAGTAATATGAAATCTTTCTACTACTCCAAAA
TAGAAAACAGAACATTAAAAAGATAAAAATTCAAACATACTTACCAGTAGATTTTC
AACTGNGCAAAGCTCATTGCATGGG

Marked-Up Copy

~~Sequence ID 873~~SEQ ID NO: 344

GTTTTCCACCGTGAAGAGAACATTTCTCTGGGAATGACAAAGCCCTCAGGAACNG
CTTTTATTTCTATTGGAAGATGCCCATCATACTTCTGGCAGGATAAAATGATAAAT
TTATTTATTCAACAGATGATACTCAATTCCCTGCTGTTTTACTAAAGGTTCTTTAC
GTTTTATAGAAGCTAAATTTACTGTCATAGAAATTGCAATTGTAGATGTTACTGTA
ATCTAGTCAGAATATCCTTATCCTTCTAAAATAAACTAGTTAAATTTATTAACAT
ACGTACTGATATTAATTTTTTAAGTTTAATGCTGCCACGTGCTTCTGCTAAGAACAT
TTATCACTACAAGTGGCAGAAAATTCCAACTCATCAAACCAAACCTGTTGCTTCT
TCCCTGCTTTTTTCAGAAAATGAGAAAGGATGACTTTATTCCAACATATTCTAAAAG
TATTCCAAGAACACTACCTTTATTCTAAATTCGTTATTTTCACAAAATAAAGGCTG
CAGATTGAAAGATAAAGGATTGCTATTAAAGAACAAAAGAAAACAAAACCGAGAGA
GAAGGAGAGCTAGGGAAATCCCTGCANAANAACCGAATANGGTCCCTCTATTCTGG
GCCGGGGCCTGAAACTATGAAACAGGCCAACACAGAATCTTGCA

~~Sequence ID 875~~SEQ ID NO: 345

CCTCTGACTCGCTCAGCTCACCCACGCTGCTGGCCCTGTGAGGGGGCAGGGAAGGG
GAGGCAGCCGGCACCCACAAGTGCCACTGCCCGAGCTGGTGCATTACAGAGAGGAG
AAACACATCTTCCCTAGAGGGTTCCCTGTANACCTAGGGAGGACCTTATCTGTGCGT
GAAACACACCAGGCTGTGGGCCTCAAGGACTTGAAAGCATCCATGTGTGGACTCAA
GTCCTTACCTCTTCCGGAGATGTAGCAAAACGCATGGAGTGTGTATTGTTCCAGT
GACACTTCANAGAGCTGGTAGTTAGTAGCATGTTGAGCCAGGCCTGGGTCTGTGTC
TCTTTTCTCTTCTCCTTAGTCTTCTCATAGCATTAATAATCTATTGGGTTCATT
ATTGGAATTAACCTGGTGCTGGATATTTTCAAATTGTATCTAGTGCAGCTGATTTT
AACAATAACTACTGTGTTCCCTGGCAATAGTGTGTTCTGATTAGAAATGACCAATAT
TATACTAAGAAAAGATACGACTTTATTTTCTGGTAGATAGAAATAAATAGCTATAT
CCATGTACTGNAGTTTTTCTTCAACATCAATGGTCATTGNAATGTTACTGATCATG
CATTGGTGAGGNGGTCTGAATGTTCTGACATTAACAATTTTCCAT

~~Sequence ID 876~~SEQ ID NO: 346

nt:

115

AAACTTTTGTGGCAACAGTGCCTAATTTGGATAATGTTTGTTCCTCAATAAATTAA
GAGCCAAATTGTAAA
AAA

~~Sequence ID 878~~SEQ ID NO: 347

nt:

634

Marked-Up Copy

GCCAGGCTTTGTGAATTACAGGACATTTGAGACAATCGTGAAACAGCAAATCAAGG
CACTGGAAGAGCCGGCTGTGGATATGCTACACACCGTGACGGATATGGTCCGGCTT
GCTTTCACAGATGTTTCGATAAAAAATTTTGAAGAGTTTTTTAACCTCCACAGAAC
CGCCAAGTCCAAAATTGAAGACATTAGAGCAGAACAAGAGAGAGAAGGTGAGAAGC
TGATCCGCCTCCACTTCCAGATGGAACAGATTGTCTACTGCCAGGACCAGGTATAC
AGGGGTGCATTGCAGAAGGTCAGAGAGAAGGAGCTGGAAGAAGAAAAGAAGAAGAA
ATCCTGGGATTTTGGGGCTTTCCAATCCAGCTCGGCAACAGACTCTTCCATGGAGG
AGATCTTTCAGCACCTGATGGCCTATCACCAGGAGGCCAGCAAGCGCATCTCCAGC
CACATCCCTTTGATCATCCAGTTCTTCATGCTCCAGACGTACGGCCAGCAGCTTCA
AAAGGCCATGCTGCAGCTCCTGCAGGGACAAGGACACCTACAGCTGGCTCCTGAAG
GAGCGGAGCGACACCAGCGACAAGCGGAAGTTNCTGAAGGAGCGGCTTGCACGGCT
GACGCAGGCTCGGCGCCG

~~Sequence ID 879~~SEQ ID NO: 348

GTTGCCGGGTCTGTGATAACTCTGTTTAACATTTTGAGGAACTGTTGAATGGTTT
TTCACAGCAGCTGCCTCATTTTTTTATTCCCATCAGCAGTACTTCTTGTTCTAATA
CCTCCACGTTCTCGCCAACACTTGTGTTGTCTGTAATTTGTTGTTAGCCATCCC
AGTGGGGATGAAGTAGTATCTTACTGTGGTTTTTCAGTTGCGTTTTCCCTGATAATTA
ATGATGGTGAACATCTTTTCATGTTCTTGTGTTGGCCATTTGTATGTCTTCTTGGGAA
AAAAAAAATGTCTGTTCAAATCCTTTACAAAGTATTTATTTTTTATGTCAACAATA
TAACCACTCAGTACACTGCTTTTTTANACAATGATCTTTTAAAGGTTTGTTTACAAC
ATTTAGCACTTGAAATTTTAAGGTTATGCCCTCAAAAAAATTGCTGAGGGAGCTAA
GCTATGAAGATGCAAAGGCATAANAATTATACAATGGACTTTGGGGGAATCCAGGG
AAAGGGTGGGAGGGGGGTGANGGA

~~Sequence ID 881~~SEQ ID NO: 349

TCGACTCTGATTTTTTTTTCTCCTTCCTCGCAGCCGCGCCAGGGAGCTCGCGGNGC
GCGGCCCCGTGCCTCCGGCCCGAGATGAATCCTGCGGCAGAAGCCGAGTTCAACAT
CCTCCTGGCCACCGACTCCTACAAGGTTACTCACTATAAACAATATCCACCCAACA
CAAGCAAAGTTTATTCTACTTTGAATGCCGTGAAAAGAAGACAGAAAACCTCCAAA
TTAAGGAAGGTGAAATATGAGGAAACAGTATTTTATGGGTTGCAGTACATTCTTAA
TAAGTACTTAAAAGGTAAAGTAGTAACCAAAGAGAAAAATCCAGGAAGCCAAAGATG
TCTACAAAGAACATTTCCAAGATGATGTCTTTAATGAAAAGGGATGGAACCTACATT
CTTGAGAAGTATGATGGGCATCTTCCAATANAAATAAAAGCTGTTCTGAGGGCTT
TGTCATTCCCAGAGGAAATGTTCTCTTCACGGTGGAAAACACAGATCCAGAGTGTT
ACTGGCTTACAAATTGGATTGAGACTATTCTTGTTCAGTCCTGGTATCCAATCACA

GTGGCCACAAATT

~~Sequence ID 883~~ SEQ ID NO: 350

TCATTTACATTAATACTCAAACTGCTCGATTAAGCAGGTGCTGTTCTTATCGCCA
TTTTGCATATGATGAGAAAGGGTAAGGTCACCCAGCTAGTATTTGGCTCACAGCAG
GCCTTAAGACTTGTTTGTGTGACTCATCAGTCCACGCTCCTAAACCACTAAGTT
GTTCTACCCTTTAATGTTGAATTAACATTGGATAGTGTTCAGTTTANATGGGTGG
GTGAGGGCCCAAGGACCTTTCAAACCTCAGATCTCTTATTTAATAACCTGGTCCCAG
ATCCATTCTCTGTCTGAAGAGGAAGTCATCCTTCAGTGGCTATTCATTGTGGGGTT
AAGAGCGCAGACTATGAATTCAGTCTTTTTGGGTCCCAGTTTGCCAGACCTTGAGT
GAGTGCCCCGAGTTTACTTACTTGTAAAGGTAGGTGGAGGTAATATAATTAATAA
ACTTAAAAAACTAATTAAAAACAAAACAAATGAACTAAGGTCTTAGGATATCTGGC
GTCTATTTTGCGCCAAATCACATAATGTCTATTGTTGTGTGTTGGACTATAGGATT
GTCCTTTAACAGGGAAGGGTTTATTTCTGTAATCAAGTCTGTCAATATTATGACCA
TGTTGATAATAGCTACCTTTAATTGAGGGCTTCCATGTGCCAA

~~Sequence ID 885~~ SEQ ID NO: 351

TCAGTGGAAAAGGGCAGGTTGAATCAAGGTGAATCAATCTGAAATTGAGCACACCT
GCCTGCCATCGCTGTTTCCTTCAACTGAGTGCTGCACATCATGGGCTCTGTCTGTGA
GAGAAAAATCCCGGTGCTTGGTGTCTTGCATGACATGGAGTTTTGCATGTAGATC
AATTTAAATGTACCTCTTGTTTACATAATTTGCATAATTTTAAAGATAATGTTG
CCAAACTTTGGAAATGTTAATGTTCANACTGAAAATCTCCACTACATGTAACCTTC
TTCCTCTGGATCAGTGGCATGGCTTATAATCCCAGCCAGTGGTTTGAAGTGTCCA
GTGTCAACTGCCATGTGCTCTGCTTCAAGGGGGAAGTAGCCTTTTGTGAATTTTTT
GTACATAAGTATTTGTTACAAATATTTTAGCAAATGCTTTCTATTTCTCTTGCTTG
TGCATATCTTGGCTGGCGTTACAGAAAAATAGTGTAACATTATTTCTTACCGGG
GAATGAGGGTTTT

~~Sequence ID 887~~ SEQ ID NO: 352

AGCACCTGGCACAGAGTAGTAGCTAACACAGATGTTAATTTGCTGCGTCAAATGT
TTTCACTTTGAATCTCTCTTGAGTATTGTTCTCCTTATTGATTACATGATGACATC
CTGTTTTCTCTCCCTGACCTTTACTGTTTGTGTTAGAAAAAAAAAAAAAAAAAAAA
AAAAAA

~~Sequence ID 889~~ SEQ ID NO: 353

CAGAGAGCTTGTTCCCTCCCTCCCTGTGCATGCAAACAAGAGGGCATGGGAGCACA

Marked-Up Copy

CAGAGAGATGGCAGCCACCTACAAGCCAAGAGGAGAAGCCTCACAAATCAAACCTCTC
GCTGCTGGCGAGAGTCTTGGACTCTGTCTTGGACTTCCAGCCTCCAGACTGTGAGA
AACAAATTTCTGTTGTTTCAGCTTCTCAGTCTCTGGTGTTTTGTTATTGACAGCCTG
AGAACACAGCTGTACNATTATNAGGGAAACAGAAAACACTGATACTTAACAATGCT
AATGCAATTATTTATTTGCTTTTCAGTCTCTACAAAACGTTCTAAAACACTAATCT
AAATATTAACAGTAAAATATTTGCATAACTAATGGAACTAAGAAATCATATGACC
AATATTTCACTTATTGGTAATCTTACTCTACTGATTTCCCCCAGACTGTGATTTT
TGAACCTCCTTGCCCTTCTCCTGTCTTCTGNGTTTATTCATGGAATTCCAGTTAT
CTGGGCTTGAAATTGCAGGCTCTCCTAACTTAAGCAAAATCTGACAGATCAGCAAA
ATGAGATAAATGTTTCTTTTTTCTTCTGACTGCATTAAATCAGATACAACCTCAGC
ATTAAAAAGCTATCTTTGNAAAATGNTGGTACTAATAAATTAGTCTTA

~~Sequence ID 890~~SEQ ID NO: 354

CCAGTTCCACATTCAGTGAAGTCATGAACTTGAAATTGGCCATGATCAAAAAGTAT
TTAAATCACAGAAGTTGCAAATGCCACAAATCAAGGTCTTTTTCTCTTGGAGAACC
TGTTAAACATTTACCAACTCACGACCGCCATGCACCCAATACTGCAATAGGTCTAT
AGATGCAGATACTGTCTCCATGAATCTTATAGGCTAGAAAGGAAATAGATAAGTAG
TCCTACCAGAAGAACATGATGAAGGCATTTGTGGTAAACAGAATGATGGCCCCCA
AAGATGTCCACATCCTAATCCCTGAAGCCTATGAATATACTACTTTACTTGGCAAA
AGGGACTTTGCCACAGGTTTTTAATTAAGGACCTTGAAATAGAGAGATTATCCTGG
ATAATCCAGATGGCCCCAGTGTAATCCCAAGGGTCCTCACAAAGGGTAGGAAGGAG
AGCCAGAGTCAGAGAAGGAGACGTAGCAATGGAGGCAGAGGTCANAGAGAGATCTG
CAGATGCTGCTGTGTTGGCTTTGAAAATGAGGAATGCAGGTGACCTCAANGNGCTA
GATGATGCAAGGAAACAAATAATCTCCTATGAACCCTAGGATGGGCATTATTATGA
GTCCTATTTTATAAACAAGGAAGTACNTCCAGAAAGATAAATGC

~~Sequence ID 891~~SEQ ID NO: 355

nt:

626

GGCAGAGGTTGCAGTGAAGTGAAGTCAATGCAATCCAGCCTGGGCAACANG
AGTGAGACTCCATCTCAAAAAAAAAAAAAAAAAAAGACAAGAGTNTCCACTCTAAACA
CTTNTATTCAACATAGTCCTGAAAGTCGTAGCCACAGCAATTTAACAAGATAAAGC
AATAAAATGTATTCAAATAGAAAAAGAGGAAGTCAAATTATCTTCACTGGNGATAT
AATTCTCTACCTGGGAAACTTCACCGAAAAAGATTTACCAAAAAGATTTCTAAGCC
TAAATAATGACTTCAGCAAAGTCTCACCATACAAAATCAACATACACAAATGAGTA
GCATTTCTGTGCACCAATAATATTCAAGCTGAGAAAAAAGAACATGGTTCTATTT
ACAATAGCTACAAACAAAAAAATATGTACCTAGTAATACATTAAATCAAGGNGGTA

Marked-Up Copy

AAATATCTNTACAACAAGAACTACAAAACCTGCTGAAAAAAAAATAGAGACACGCAAA
TAAGTAAAAAGGCACTCCATGCTCATGAATTTAAAGAATCAATATAATTAAAATGT
CCGNGCTGCCTAAAGCAACTTACAGATTAAAGGCTATTTCTCTCAAACCTATAAATG
CACCTTTTTTA

~~Sequence ID - 893~~ SEQ ID NO: 356

nt:

585

GTCATTGCTGGGTGGCGCCAGCCCTCAGACTTGCCTCTTTGCAGTAGGAAGAAGGC
CTCCCCACATACCTTCCCACACTCATCACCTTAAGCCAGACTCGGTGTCCAGTGAA
TATGACCATCTCTTGCCCATTTTCTAATGAGTGTTTTCATTAATGAGTTATAAGAA
TGTGGTGGGTAAATCTATGGGCTTTGAACTAGTGAATCAACTTGGTTTTCAGAATCT
GGCACTGCTACTTACTAGTGAATTTAAGCAAGTTATTTACCTTTTCAGAGTGTCAG
TTCCCTCATGCATACAAGGAAGATAAAAAATAATGTNTACNAAAGTATTGGAGTAA
TTAATACATGGAGAACTACATGTAAAGCGTTTAGCATGATGTCTGACATATTAAGC
ATCCAATATTAGTNGCTTGCAGAATTATTAGTAAAAGAGATTGCTTCTGAAAGCCA
TTCCAATTCTTAAATTTTATAATGCCACATTTGAGGTCACCTGAAGTCGTGTATAA
CATGTGTACATTTTTTGCGATTTATTTTTTCAATTCCCANATTAAAGGCATAGAGAT
ATCCTAGCNANGGACTCCAAGTGTG

~~Sequence ID - 895~~ SEQ ID NO: 357

nt:

560

GTAATTGCAGCCTGGGCAACGGAGTGAGAGACTGTCTCAGGAAAAAAAAAAGAAAA
AAAACCTACTGAGGTAGTTGAATATATCCTCCATTCCTCCATTTGTGGATTAGTTAGT
AAATGGGGCATCTTAGGGTTTAAATATGTCCAGGGTCACTGAGGATCAGATCCTAG
GGTTCCTTTGACTCAAGGCTTTTGTCTCAGCAAAACGTCACCTTCCAGCAGGAAGG
CTTTCTCAGGCAAGTAGCAGGGTGGCTACTATGTATCGCTTCTTTATTTTTTCTTT
TTTAAATAATGCAGGCACCGTGCGCATAATTTAAAAAATCAGTGCTAAAACCTT
AAAAAAAAAAGCTGTTCTCATCTCCTGTCTTTCTTTTTTTTTCTTTTTATTTTT
TTCTTTTATTATTATTATACTTTAAGTTTTAGGGTACATGTGCACAACGTGCAGGT
TTGTTACATATGTATACATGTGCCATGTNGGTGAGCTGCACCCATTAACTCGTCAT
TTAGCATTAGGTATATCTCCTAATGCTATCCCTCCCCCTCCCCCTTTTTTTTTTT

~~Sequence ID - 896~~ SEQ ID NO: 358

GGGAATGTCTTAGGCACTGGGACTGTAAGTGCAAAGACCCTGTGGCACAAGGGAAT
GTTAATTATCTACCTTTCANAACTGGAANAAGGCCTAGCCTAGAGCATTGAAAAC
AATAAGGGAAAGGAGGAGTAAGGCTGGANAGATAGGAATGGTTTAAAGTCTTTGTT

Marked-Up Copy

AAAAATTTTTTTTAAAAAATCTTTATCACAAGAAGAGGATTGGCNTGATCAAATTT
GACTTTTAAAAANATTACTTGGGTGGGCATGATCAAATACTACTTAGGGAGATTA
GTTTANATGATAATGGCATTCTGGACCANAGTGGAGTCAGAGGTGAAAAGAGGTAG
ATATTCCANAATTGAGGGATTTGTGAGGTGAAATCATTGTGTTACAGATATTAAAGG
ATAAGGAGCTTTGTCAAAGGGGATCTTAAGTTTCTGGTATGGTAACTGGGTAGAG
AGCCCTGGAACATGACCAGCTTTAAGGGAAGAGAGCTTGAGCTCTGTTCTTGTTAA
GCTCAGTTTGAGATCTTTGTGGAATCAAGTGGAGAGGTCTAAGCAGGGAAGTGGCT
TGGCTAGGCTGTAAAGATGAATCTGAGAGTCCCAAGAATATGGTAATTATTAATAA
AAGCCTTAGGTANATGAAATTGTTTTGGG

~~Sequence ID 897~~ SEQ ID NO: 359

nt:

509

GCAAATCTACACATTTGATTAAATGATAGGGAAGTATGCACACACATAATACATAT
AATGCTAGTTTCTTGGTTTTGATATTGTACCATAGTTATGTAAGATGTAACCATTG
GGGGAAGTGGGTGAAGGCTACATGAGACCTCTCTGTACTTAATCTTTGCAACTTA
TGTGAATCTATAATTATTCAAAATAAAAAGTTTTAAAGAACCTAAGTATCCTTAT
TACTGAGGGTCATCGTGCTAGACAGCAAGGTTGGGCCAGAGCTTCTAGTTATTTAA
AATACTAAATACCAGCCTGGGCAACATAGCAAGACCCTGCCTCTACAAAAGCAAA
AAAATTAGCTGGGCATGGTGGTACATGCCTGTGGTCCTAGTTACTCTTGGAGGAGT
CTGAGGTGGGGAGCTTGAGCCTAGGAGTTTGAGGCCGCAGTGAGCCTTGATTGTGT
CTCTGTACTCCAGTCTGGGCCACAGAGCAAGACCCGGTCTCTAAAAATAAATAAAT
AAATA

~~Sequence ID 898~~ SEQ ID NO: 360

ANTGCACTCCAGCTTGGTGACAGAGGGAGACTCCATNTTAAAAAAAAAAAAAAAAAA
AAAAAAGGGAGTAGCTTGAAGCCACATAGTAGTTAGTGGTAAAGGCCACCCCTTT
TCCCACAACCTCACACCAGCACCAAGCTAGCCTTTNTAATTTCCAAGCCAGTGCC
CTTTCAACGCACACACCCCTGTGTCAGTTCCTTTCTGCTGCAAGCTCTCTGGAGG
CAGATACTGTTGAGTCCCTGGCCTGCCTATGAGAACGGCTCATGATCTCTATTTCT
TCTGCTTAATGACCATCTCGAAGTAACAAGTTTAGCCTAAAATAAACTTGCTAAGT
TAGCAAAGGAAGTCCTTAGCAGCCACCATTTCTCGATTCCCTCCATCACCTCCCCTG
CCCCTCAACTCCCTCATTTCTCCAAGATATGGGCTCCAGGCTGGGCGCGGTGGCT
CACGCCTATAATCCTAGCACTTTGGGAGGCTGAGGTGAGCAGATCACTGAGGTCAG
GAGTTCG

~~Sequence ID 899~~ SEQ ID NO: 361

TCNTTCGGAACGCGCC

~~Sequence ID 900~~ SEQ ID NO: 362

CTGGAGGGATGGGTAGGATTTTGACAAGAGTGGTTGAAGGTATTCTAATTCACCTTA
GTACCTACATGTGCGAGGCAGCATGAAGGCCAAAAAGCCTGGGGCATG TTCAGAGA
ATAGCAAGTATTCTAGTTTGAGTGGCACCTGGTACGTATATAAGGGAATAGTAAAA
GATCTGGCTGGAAAGGAAAAGTAGGGGCAGGTTACGAAGGACCTCTGAAAGTCAGA
CTGTGGAACCTGGAAC TTTTATCAGGAAGCAGTAGTTAGTTTTTTTCAAGCAAAAGCT
AATTAGAGTTGATATTTAGGAGGATGAATCTAACAGTTGTGTGCAAGGATGCCTTC
AAACTGAGTGAGACTAGTACTGGAGACTGGTTAAGAGACTACAACAATAACCTGAG
TAAGAATTAATACAGGCCTGACCTAGTTTTTGAGTGAGTAGGATTGGAAACAAGAGT
TTTAGGTATTATAGGATTTATGCATATAAAATGGACTTGACAGAACTTGAAGAAAG
AGAAAGTGTCAAAAGGACACAGAAAGTGAGGCAGGATATCTTACAATGTTAAAGGA
AAGGAATAATAGAAGTTAC

~~Sequence ID 903~~ SEQ ID NO: 363

GGAAACATAAGCTTGTTTCAGTACACTCACGCTGTAGATTAATTCTGATATTACAT
ATCTCCATCAGACTTTGTACCCTCTCTCTTCCATCCCTTACCCTTACCGATTAGGT
TGGTATTACCTAAAAATCCATAGAAAATGTCCAGGTGAATTGCCTTATGCTTTCTA
CCCCATAAGGTATAATT

~~Sequence ID 904~~ SEQ ID NO: 364

CTCTGTGGTGTGAGAACACAGTGGGTGACCAAGGCTTTCAGATGAACCCAAGGAA
AGTGAAAAAGCTGATGCTAATAACCAGACAACAGAACCTCAGCTTAAGAAAGGCAG
CCAAGTGGAGGCACTCTTCAGTTATGAGGCTACCCAACCAGAGGACCTGGAGTTTC
AGGAAGGGGATATAATCCTGGTGTTATCAAAGGTGAATGAAGAATGGCTGGAAGGG
GAGTGCAAAGGGAAGGTGGGCATTTTCCCCAAAGTTTTTGTTGAAGACTGCGCAAC
TACAGATTTGGAAAGCACTCGGAGAGAAGTCTAGGATGTTTCACAACTACAAAGC
TGAAGAAAATGAAGCCCTATTACTTGTTTGTAAGATTTAGCACCCCTTCTGCTGTAT
ACTGTACTGAGACATTACAGTTTGGAAGTGTTAACTATTTATTCCTGT TAAATTT
TAACCTACTAGACAATGATGTGAGTACCCAGGATGATTTCCCTGGGGCACAGTGGGT
GAGGAGATGGGGACAGGTGAATGGAGGAGTTAGGGGAGAGGAAAAGTGGAAGG
TGTCTGGAAAGGGCACCAAAAAAGTCTTCCAGGTCTGATCCTGTTTCTTGCTCTGA
GTGCTAGCTACCACTGTGTCACACTGTAACATN

~~Sequence ID - 905~~ SEQ ID NO: 365

nt:

655

CTCAGCTCTTGCCTGGTCACCTTGTGGCTTTTACCATCCTCATCCCCTGTGCCACC
CACATCCTGCCACTTCTGCATGGAGTTGGGGTGGGGCCATTGGAGAAAAGAGGTTA
AACAAGCAGTAATTTACTTGAGTACAGTCTTTGAGCCAATGAAATGCCAGTCATCA
TTTCCCAGGGGTACTTGTTCATCTTGTCAACAACCCGCTGATAATGCTCCTTCAATG
TGAATAGCAAAAGTAGGGAGAGACGCTGAATGAAGAAGATGCCTACCCCTCAGGAA
GACTGCTGTCCGCCTCCAGGCCTGCATGCACACACCCATGCCACCTGCACCCCCA
GCACCACGCCCACACTCACTCGCACACACCCACATGCCAGTGTTTTGGGGTTGGCA
GCCTGGACACTGCTGAGGCAAACACAAGTCATCAAGCATAATTCTCATTCTCTCCT
TCTGTCTCTGTTTTAGTTACAGGAATTTGGTCAGTTTAGAGGATTTAATAAGTCCG
TGGAAAATTTGTTTCTGTCTCTTGCTACCCACGTGAAAAGTAAGTGCATGCTTCAT
GATGTGTTTTCCCACTACCTTCCAGGCCAGCCGAGCCCCTGGCCANGGCCTGGCC
CGGTGACCTCGGTTGACACTGTCCTCANGCCACTCACTT

~~Sequence ID - 906~~ SEQ ID NO: 366

CAGAATTTTCATGTTTATGCTGCACAAGGCCTGTATTTTATAATGGTGGCTCTTTTG
GACGATGACTTCCTCGATGGTGAACTTCCAGTAATCTCCCTCATCATACTGAAAT
GATATCAGTATATCATCAGAACACCATGGAGCTTGTTCATTTGAGGGACACAGCTTG
CTTGTGTGCTTGGGAAAGAAGAGGTTTAGCATGGTTTCAGGTCAGTGATGAGTCCA
ATGATCTCTGCAAGTTCCCTTAGCTCTGANAATTCTGATGTCATATGCACTTCTGC
CGCCAGAGTTGCTGCTTACTGGATGCGTAAGAAGAAAAGAAAAAAAAAAAAAAAAA

~~Sequence ID - 907~~ SEQ ID NO: 367

nt:

582

CTTCCATTGGGGGTAAAGATCAAACTTTAGGCGAGCCAGGTCTGTATCTCCATTCC
TGTCTCTGACTGCTTCCCTGTAGGGATTGTCTGCAAGCGCACACCTGCATTTTCTT
GTCCACAAGTCTATGCTCTAACTCTGTACCTGCATGGCTGCAAATTAGCTTCCTT
CTTCCTGCCCTCTTCTCTCTAGCTTGGATTTTGAATTTGAATGGCAGGCATGGGAT
GTCCGTGTGTGTGTACTGCTGATGTGTACAGCCGCTTGTTAGCGCTCTCATTTGTCT
TCAAATGTAAGTCATTTTGGCTGGGTGCGGTGGCTCATGCGTATAATCCCACGCTT
TGGGAGGCTGAGGTGAGCTGATCATTTGAGGTTAGGAGTTCGAGACCAGCCTGGCC
AACATGGCAAACTCCATCTCTACCAAAAATACAAAATTAGCTGGGTATGGTAGT
GCACGCCTGTAATCCCAGCTACTTGGAATGCTGAAGCAGGAGAATTGCCTGAACCC
ANGAGGCGGAGGTTGCGGTGAGCCAAGATCACGCCACTGCACTCCAACCTGGGTGA
CAGAGCAAGGCTGTGTCTCAA

~~Sequence ID 908~~ SEQ ID NO: 368

ACCTGACTTCAAACCTATACTACGAGGCTACAGTAATCAAAACAGCATGGTACTAGT
ACAAAAACAGACCAATGGAACAGAATAGAGATCTCAGAAATAAACTGCACATCTA
CAACCATCTGATCTTCAACAAACCTGACAAAACGAGCAATGGGGAAAGGATTCCCT
ATTTAATAAATGGTGCTGGGAGAACTGGCT-AGCCATGTGCAGAAAATTGAACTG
GACCCCTTCCTTACACCTTATACAAAATTAACCTCAAGATGGATTAAAGACTTAAA
TGTAGAACCCAAAACGATAAAAACCCTAGAAGAAAATCTAGGCAATATCATTAAGG
ACATAGACATGGGCAAAAATTTTCATGATGAAAACATCAAAAGCAATGGCAACAAAA
GCAGAACTGACAAATGGGCTTCTGCACAGCAAAAAGAACTATCGTCAGAGTGAAC
AGACAACCTACAGAATGGGAGACAGTTTTTGCATCTATCCATCTGACAAAAGTCT
AATATCCAGAATCTACAAGGAATTTAA

~~Sequence ID 910~~ SEQ ID NO: 369

CAAAAAACAAGAATTACCCGGGCTTGGTGGTGCATGTCTGTAGTCCTATCTACTCA
GGAGGCTGAGGCTGAAGGATCACTTGAGCCCAGGAGTTTGAGGCTGCAGTGAGTGA
GCCATGATCATGCCAGTGTACTCCAGCCTTGGCAGACTGAGCAAACTTGGTCCCT
CGCAAAATGTTGAAGCCCAGTTTTTCACTATTAACCTGTATTTTCAGTTTCCCCATGC
TAACTTTGAAACACTGGGGCTGGCCTGAGGGTATAAAGGCTTATTCAAACCTCAGTA
ATTTAACTTAAAATCCTAAGGAACCTCAAAAAGTGTAATCTAGTCCAAATGGGGC
ATCAATTCTAAAGCATTTGCTTGTTTGAGCAGATTTTCTGTGTCTGAGGTATATAG
ATAACTTATCTTTTTATGACTAAATCCAAGTCCTTAGTTCCTGTTGGAATTCAAAA
TCATATTTAAAAATTGATGCTTTGTTCTATAATTAATGCTTTGATTGTATAAATAA
TAAGTATTCTTCAAATCCCTTTTTACAGATGATGATTCTGATACCGAGACGTCAA
ATGACTTGCCAAAATTTGCAGATGGAATCAAGGCCNGAAACAGAAATCAGAACTAC
CTGGNTCCCAGTCCTGTNCTTAAAATTCTAACTCGAC

~~Sequence ID 911~~ SEQ ID NO: 370

nt:

595

GAGGGTGTAGAAGAGAAGAAGAAGGAGGTTCCCTGCTGTGCCANAAACCCTTAAGAA
AAAGCGAAGGAATTTTCGCAGAGCTGAAGATCAAGCGCCTGAGAAAGAAGTTTGCCC
AAAAGATGCTTCGAAAGGCAAGGAGGAAGCTTATCTATGAAAAANCAAAGCACTAT
CACAAGGAATATAGGCAGATGTACAAANCTGAAATTCGAATGGCGAGGATGGCAAG
AAAAGCTGGCAACTTCTATGTACCTGCAGAACCCAAATTGGCGTTTGTTCATCAGAA
TCAGAGGTATCAATGGAGTGAGCCCAAAGGTTTCGAAAGGTGTTGCAGCTTCTTCGC
CTTCGTCAAATCTTCAATGGAACCTTTGTGAAGCTCAACAAGGCTTCGATTAAACAT

Marked-Up Copy

GCTGAGGATTGTAGAGCCATATATTGCATGGGGGTACCCCAATCTGAAGTCAGTAA
ATGAACTAATCTACAAGCGTGGTTATGGCAAAATCAATAAGAAGCGAATTGCTTTG
ACAGATAACGCTTTGATTGCTCGATCTCTTGGTAAATACNGCATCATCTGCATGGA
GGATTTGATTCATGAGATCTATACTGTTGGAAAAC

~~Sequence ID 912~~ SEQ ID NO: 371

nt:

651

CATTTCCAGAGTTTATGTGAATTGAATTGAACTATGGTTTTATGTTACTGTCAGTA
GAATGAAGTACGAATATTTGAAAAATACACCTTCAACTTCAAAGTGATTCTTGACA
AAAATTATAAGGAATCATTTTGGACACATTTTCTGGTAGAGCCTTGTA AAAATTAA
AACCAAGTGTTGTTTTCAAGAAGAACTGTAATACATAATCAGGAATTTGAGTAGGG
AGATTATTTTGTATTTTAAAATTAAAGTGGCTGTGTAGTTTTAACTTTAGTATTGC
AGGTAGAGTAAGCTTACATGATAACAAAAATCTTGGTCTTAGTGACTTAATGATTC
TGATATTTATTGATTGATTGGTTATCATTCCAAATATTTTAAAAGATAATAGCTGG
CTGGGTGCGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAGGACGGGCG
GATCACGAGGTCAGGAGATCAAGACCATCCTGGCTAACACGGTGAAACCCCGTCTC
TACTAAAAATCAAAAAATTAGCCGGGTGTAGTGGCGGGCACCTGTAGTCCCAGCTA
CTCAGGAGGCTGAGGCAGGAGAATGGCATGAACCTGGGAGGCGGAGCTTGCAGTGA
GCTGAAATCGTGCCACTGCCTCCACCTGGCGACAA

~~Sequence ID 913~~ SEQ ID NO: 372

GTGAGGTGGGGACTTCATTCATTGTCCTATTTCTATCTCCACTTTGTGCCTGGAGA
GCTTTCAGGGGAGGTGGAGGAGGAGGGTCTGCCAAGCTACTGCAACATCTGTCACC
CACTATAACCAGTTACTTGGGGGAGGACAGACACTGTGGTGTCATTAAAGTTGTTT
GAACCAAAGTGGCGGCTGCATCTTTGTCCCGATGCTAGCCGTGCCGGTCTCCCATC
ATCCGCTCGCCCTCCTTTNCCCTGGGCTGCGCCCACTTGTCTTCCTGGATATTTGG
GGGTGACTCGCCATGCTTGGCACCCCTCTGCTTCCTGGTGCTGCTCTGACTCGAAGA
CGGGACAGTCCCTGGTGCACATCCAGGGAAGAGGAGTGTGCGGTAGTTCTTGACAGTA
GGCACTTTATCAGGACCTGACCTGTTGCTGGGTGATTTTAGTCTCTACAAACAGAA
AGCGTTTCAAAGCGTCAGCTGTGGGAGCAGAGTGACCCTTTGCTGATGCTGGGGGG
AGGGGATCTAAATCCTCATTTATCTCT

~~Sequence ID 914~~ SEQ ID NO: 373

GGCGCCTGCTGGAGGAGGAGAGCTCTGCTGGCATGAGCCACAGTTTCTTGACTG
GAGGCCATCAACCCTCTTGGTTGAGGCCTTGTTCTGAGCCCTGACATGTGCTTGGG
CACTGGTGGGCCTGGGCTTCTGAGGTGGCCTCCTGCCCTGATCAGGGACCCCTCCCC

Marked-Up Copy

GCTTTCCTGGGCCTCTCAGTTGAACAAAGCAGCAAAACAAAGGCAGTTTTATATGA
AAGATTANAAGCCTGGAATAATCAGGCTTTTTTAAATGATGTAATTCCCAGTGTAA
AGCATAGGGATTTTGGAAGCAGCTGCTGGTGGCTTGGGACATCAGTGGGGCCAAGG
GTTCTCTGTCCCTGGTTCAACTGTGATTTGGCTTTCCCGTGTCTTTCCTGGTGATG
CCTTGTTTGGGGTTCTGTGGGTTTGGGTGGGAAGAGGGCCATCTGCCTGAATGTAA
CCTGCTAGCTCTCCGAAGCCCTGCGGGCCTGCTTGTGTGAACCGTGTGGACAGTGG
TGGCCGCGCTGTGCCTGCTCGTGTTCCTACATGTCCCTGGCTGTTGAGGCGCTGC
TTTAACCTGCACCCCTNCCTTG-CTCATANATGCTCCTTTTGA

~~Sequence ID - 915~~ SEQ ID NO: 374

nt:

230

TTTGAGACCAGCCTAGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAAA
TTAGCCGGGCGTGGCGGCACATGCCTATAATCCCAGTACTTGGGAGGCTGANGTA
GGAGAATCGCTTGAACCCANANAGGCAGAGTTTGCAGTGAGCCGAGATTGTGCCAT
TGCACTCCAGCCTGGGCGACAGAGCGAGACTCCATCTAAAANAAAAATAAATGAATA
AAATAA

~~Sequence ID - 917~~ SEQ ID NO: 375

NNCAGATTTTTTTTTTTTTTTCAGNGTTAGACCATCTTCAATTCCTGGAACAAAC
TTAACTTTCCATGATATGTATTTTTTATACATTGCTGGATTTTATTTGCTAATATT
TTACTTAGGATTTAATTTTCTAAGTNGACCTATAATTNTCCTGTATAAAATTGCAT
TTGTCACATTTTAGTATCAAGGTTGTCCTANCNCCATGAAATGGATTTANAATGGT
TTATGTAANATAAAGTACATTTCTTCTAAAGGTTTGNGTGGATTAACTTTCAAATC
TGCCANAGNGNGTTTTTTTTTCTTTTTTTTTTTTTTTTTCATTTNAAGGGAGNGCAAGT
ANCTTTTCAAATNCTGATTTAATTTTAAAATATTTNCAAGTNTNTTANAGTTTT
TATTTNTNTNGAANGTTAACATTTTATANAAAANGGTNTTATCTTTTTTAAATTC
TTTGACATCAGTTTCTTCANAATTCCTTCTTTTAA

~~Sequence ID - 926~~ SEQ ID NO: 376

GTCATATCTCTTCCCAGGGAAAGCAGGAGCCCTTCTGGAGCCCTTCAGCAGGGTCA
GGGCCCCCTCGTCTTCCCCTCCTTTCCCAGAGCCATCTTCCCAGTCCACCATCCCCA
TCGTGGGCATTGTTGCTGGCCTGGCTGTCCTAGCAGTTGTGGTCATCGGAGCTGTG
GTCGCTACTGTGATGTGTAGGAGGAAGAGCTCAGGTAGGGAAGGGGTGAGGGGTGG
GGTCTGGGTTTTCTTGTCCCAGTGGGGGTTTCAAGCCCCAGGTAGAAGTGTTCCT
GCCTCATTACTGGGAAGCAGCATCCACACAGGGGCTAACGCAGCCTGGGACCCTGT
GTGCCAGCACTTACTCTTTTGTGCAGCACATGTGACAATGAAGGACGGATGTATCA

Marked-Up Copy

CCTTGATGGTTGTGGTGTGGGGTCCTGATTTTCAGCATTCATGAGTCAGGGGAAGG
TCCCTGCTAAGGACAGACCTTAGGAGGGCAGTTGGTCCAGGACCCACACTTGCTTT
CCTCGTGTTTCCTGATCCTGCCTTGGGTCTGTAG

~~Sequence ID 938~~ SEQ ID NO: 377

TGGCCATCCTTTTCCCCCAAACACACCCCCTTAACCTATCTCTTGGGACTTAGCC
CGACCCTCCCTCTCATTTCCCATTAAGTCTGAGAGGCAAGAGCTAGGTTAGGCAAG
GAGGTGGTTGGCCAGAGATGGGGAACAGCCAGGTGCCCCAGTCCTCTGATTTTTTC
TCCATCCTGCTTACCACCTCCCTGGGTACTTACAGCCTTCTCTTGGGAACAGCCGG
GGCCAGGACTGGGTACCTATGAGCTGAATCAGCATCTCCTCCTGAGTCCCAGGGC
CCCTGCAGTTCCCAGTCTCTTCTGTCCTGCAGCCCTTGCCTCTTTCCACAGGTTT
CACTTTATATCCACCTTTTCCTTTTGTTCATTTTTATTTTTATTTTTTTATTAT
TAAATGATGTGGTCTATGGAAAAAAAAATAAAAATCTGACTTAGTTTT

~~Sequence ID 939~~ SEQ ID NO: 378

nt:

513

GGAACCCAGTGTATTACCTGCTGGAACCAAGGAACTAACAATGTAGGTTACTAGT
GAATACCCCAATGGTTTCTCCAATTATGCCCATGCCACCAAAACAATAAAACAAAA
TTCTCTAACACTGCAAAGAGTGAGCCATGCCTGTAAACACTGTAAAGAATGTAACA
TGTGGGGGACACACAGGGGCAGATGGGATGGTTTAGTTTTAGGATTTTATTAGTGCA
TGCCCTACCCTCTGGGGGAACGTCCCATCTGAGGTTTTCTTCTCGGTGGGGGGATT
TAACTTCTGTCTAGGGAAAACAGTGTCTGATGAGGAGTGTTTCCAACACAGGCTA
CATGAATTTCCCTATACCAGTGCGAAAGCAGCCAGGAGTCCCCGTTGGAAAAGAAC
AATGCCACTCTCTTTTATGTATCTTGGTTCTGCAACTCATTTGTTGTAAGTAGGGT
TAATCGAGTATCAGGTTACAGTATCCTGCCCTTATTATTTTATGATTCACTGACT
CAAGTTCCA

~~Sequence ID 947~~ SEQ ID NO: 379

GAGAGTGAAAAAATTCTGGTACAAATTGGGAAATTAGTATATAACAACATAGTGTT
AAATTCAATGGGAAAAGTTTAATAAGAGGATTTGGTATCAACTGGCTGTCCAAAGA
TAAAAATGGACCGTCCTATCACATACAAAATTGTTTTTTAGATAAAGATTTAAATA
CAGGCACTCCTTCATTTGCGTGGTGCACCTTGAGGTGTTGCAGAAATGATGAGAGC
TGAAACTGCAAAGCAATTTTAATACTTTATCTGTTGGAAATCTTATAGTTTTCTCTG
TGACCGTTAAAATTTTCATTAAACTATTAAAAACACCCATGACTGGTCACAAATGT
ATTGGGAAATGGAAAAGAATTAATACACTAAAAATACAAAAAATAGAAAATATTTA
AAATTATCTAAAAATTTGAAACATTAGAAAAATTGAGAACTAGGCAGGGCGTGGTG

Marked-Up Copy

GCTCACATCTGTAATTTTAGCCCTTTGGGAGGCTGANGCAGGTGGATCACCTGANG
TCAGGAGTTCGAGACCAGCCTGCCAACGTGGGGAAACCCCGTCTCTACTGAAAATA
CAAAAATTANCCGGGCATGGTGGCACAAGCCTGTAATNCTTGCTNACCAGGANGCT
GAGGCAGGAGAATCACTTGAACCCANGANG

~~Sequence ID 949~~ SEQ ID NO: 380

GTTTCACATGAGAAGGTAGTATTATGTACAGTGACCTTGTTTAAAGTGTCNGTTTA
ATGTTACCACTAAGGCCCTGCCCCAGCTTTATCACCTGAGCACTAACAAGTGCTGT
GTGGAGTTCAGTCCATGCTGGTAACTNTTGAGTATTCAGTGGGTCTTTTAACAATT
ACCACCGTGGAGGANANAGCAAGGAAGAGAAATGCTGTGATCTTTTNCTGTTTTTA
ATTAGNGAAAGAGGGATTANATTAAACAAATGTTACAGAGNTGTGACTNTGATCCC
CCAGNGGTAAGCAATAATTGTANAGACTGGATTTNANAAGCCCTGAGAGTTTATTT
TCAACCTATNTATTATAGNNCAATCC

~~Sequence ID 1028~~ SEQ ID NO: 381

ACAAGGCTTGGGGGCTGGACTCCCTCTACTGCCTCTGGCCATACCCCTCCTGGAG
ATGGGGTCAAGGCACCAGGACTGA

~~Sequence ID 1056~~ SEQ ID NO: 382

nt:

435

TCGCTTGTAAGCCTGAGACAGCTGCCTGTGTGGGACTGAGATGCAGGATTTCTTC
ACACCTCTCCTTTGTGACTTCAAGAGCCTCTGGCATCTCTTTCTGCAAAGGCATCT
GAATGTGTCTGCGTTCCTGTTAGCATAATGTGAGGAGGTGGAGAGACAGCCCACCC
CCGTGTCCACCGTGACCCCTGTCCCCACACTGACCTGTGTTCCCTCCCCGATCATC
TTTCCCTGTTCCAGAGAAGTGGGCTGGATGTCTCCATCTCTGTCTCAACTTCATGGT
GCGCTGAGCTGCAACTTCTTACTTCCCTAATGAAGTTAAGAACCTGAATATAAATT
TGTTTTCTCAAATATTTGCTATGAAGGGTTGATGGATTAATTAAATAAGTCAATTC
CTGGAAGTTGAGAGAGCAAATAAAGACCTGAGAACCTTCCAGA

~~Sequence ID 1071~~ SEQ ID NO: 383

NGATATAGTNCCGCATGGGAAAGATGANCAGGTATAACCNAGCNTNATATAGCAAG
GACTAACCCCCCTGCCTTCTGCATAATGAATTAAGTAGAAATAACTTNGCAAGGAG
AGCCAAAGCTAAGACCCCNAAACCAGACGAGCTACCTAAGAACAGNTAAAAGAGC
ACACCCGTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGGCGACAAACCT
ACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAACTTTAAATT
NGCCACAGAACCCTCTAAATCCCCCTTGTAATTTAACTGTTAGTCCAAAGAGGAA

Marked-Up Copy

CAGCTCTTTGGACACTAGGAAAAACCTTGTAGAGAGAGTAAAAAATTTAACACCC
ATAGTAGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACACCCACT
ACCTAAAAAATCCCAAACATATAACTGAACTCCTNACACCCAATTGGACCAATCTA
TCACCCTATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCCTCCGC
ATAAGCCTGCN

~~Sequence ID - 1074~~SEQ ID NO: 384

nt:

689

GGGAGGCGGAGGCTGCAGTGAGCTGAGATCGTGCCACTTCATTCCAGCCTGGGCAA
CAAAGCGAAACTCTGTCTCAAAAAAAAAAAAAAAAAAATTTGTTGACTGTTGTAA
TTTAAAGCTTGTCAATTTTTTTATTTAGTAATAACACTCATTAGTGTAGTATCTATGA
TGAACCAGGTTCTGCACAAAGTACCTTATGTTTCATGGCCTCATATCGTCTTCTCCA
AAACTCTGCAAGATAGGATTCATCACCACCTTATAGGGAGAGATCTGAAAGTTTAAA
ATTGTACCCAAGGTCACACAGCTGGTAAGTGCCAGAGCTGGGATTCCGTAGGGTGT
TCANAGTGCCTCTCCTGCCGTAGGCTTATCACAAAAGTCAAAGTTTGGTCATAAT
AAAGCCTGAAGTTTGGCAGGATTTAAAAATAGTCACCANACTTTTGAGTTGGAGCA
TCCCACCTCACTGCTGTTTACCTTCTGTGGCAGGGAGAGTCATCATTTCCATTTCA
GCTTGTGGAATATCTTGTCAATTAACATTCTCATGCAAAGCCATTTTATGGTGCCC
AATGAANATGGTTAAGCTACTGCCCCAAGCCTNTGGAAGCCTTCCTAATTTTGGAC
TTGCACTATGCAAATTGNATAATATTTTCTCTACCCTAAGCCAAATATTTTCTTCA
CTTTTCATTCTCTAC

~~Sequence ID - 1081~~SEQ ID NO: 385

CGCCGCCGCGCCGCCGCTCGCTCTCCAACGCCAGCGCCGCCCTCTCGCTCGCCGAGCT
CCAGCCGAAGGAGAAGGGGGGTAAGTAAGGAGGTCTCTGTACCATGGCTCGTACAA
AGCAGACTGCCCCGAAATCGACCGGTGGTAAAGCACCCAGGAAGCAACTGGCTACA
AAAGCCGCTCGCAAGAGTGCGCCCTCTACTGGAGGGGTGAAGAAACCTCATCGTTA
CAGGCCTGGTACTGTGGCGCTCCGTGAAATTAGACGTTATCAGAAGTCCACTGAAC
TTCTGATTCGCAAACCTCCCTTCCAGCGTCTGGTGCGAGAAATTGCTCAGGACTTT
AAAACAGATCTGCGCTTCCAGAGCGCANCTATCGGTGCTTTGCAGGAGGCAAGTGA
GGCCTATCTGGTTGGCCTTTTTGAAGACACCAACCTGTGTGCTATCCATGCCAAAC
GTGTAACAATTATGCCAAAAGACATCCAGCTAGCACGCCGCATACGTGGAGAACGT
GCTTAAGAATCCACTATGATGGGAAACATTTTCATTCTC

~~Sequence ID - 1083~~SEQ ID NO: 386

nt:

198

Marked-Up Copy

GCGCGTCGACTTTGTTTAGACATTGAATGACTTTGTTAAAGGCACAATTAATCACA
TTGGTTGTACTCTGNNGACAGCCTTCTTTAAAAAAAATAAACAATTTAAAACAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAANTTTTAACC

~~Sequence ID - 1084~~ SEQ ID NO: 387

nt:

198

GCGCGTCGACTTTGTTTAGACATTGAATGACTTTGTTAAAGGCACAATTAATCACA
TTGGTTGTACTCTGNNGACAGCCTTCTTTAAAAAAAATAAACAATTTAAAACAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAANTTTTAACC

~~Sequence ID - 1099~~ SEQ ID NO: 388

nt:

561

TGCATGCTTGTGGATTGGAAAACTTTGGAGACTGATTACTTTTCATTATATATGT
GTCACAGTGAAACAGCTTTTATGTGTCATGTAAGATTACTGCTTGCCTCTCTAAGG
AAGGTCGTGACTGTTTAAATAGACGGGCAAGGTGGAACCTTTTGAAAGATGAGCTT
TTGAATATAAGTTGTCTGCTAGATCATGGTTTGTATTGAACTAACAAGGTTTGCAG
ATCTGCTGACTTATATAAAGCTTTTTGATTCCTACTAAGCTTTAAGATTTAAAAAA
TGTTCAATGTTGAAATTTCTGTGGGGCTCTATTTTTGCTTTGGCTTTCTGGTGAGA
GAGTGAGGAAGCATTCTTTCCTTCACTAAGTTTGTCTTCTTGTCTTCTGGATAGA
TTGATTTTAAGAGACTAAGGGAATTTACAACTAAAGATTTTAGTCATCTGGTGGA
AAAGGAGACTTTAAGATTGTTTAGGGCTGGGCGGGGTGACTCACATCTGTAATCCC
AGCACTTTGGGAGGCCAAGGCAGGCAGAACACTTGAAGGAGTTCAAGACCAGCGTG
G

~~Sequence ID - 1109~~ SEQ ID NO: 389

TTTGNCGGTNTTGGANNNNNANAANTTTCTTCCANNCNTNACNTNTTGGTGGNCTA
AATTAANATGGNTTTNGNGGGTTCNTTNCTNNNTNNNNCATGGGANANAATTNATT
NTCNTNCNNNTTCCTTNNCCCTNAANCTACCTTCCCCCNATTTTCTCCCCTNTTCN
TNAATTANCATCCTCTCCNCNTANNTCNANACNTTAATGGCAANACTATCTAATAN
CNANNATAANANCTCCTGTNNNCCACATNTCTTATTNNNCGCNNCANGTTNCANNC
CCNCAGAGTNAACTCATCCTCNCNNAANTTCATATCGTGNNCTNTNNNCNNTNGC
GCGANATATTAANNANACCNGTANNNTNNNANACANNANNTNNGNAANAANCCTTCT
NANNTTTTAGCNTCNNGCNNTAACNNNNNTCTTNGTGNNNNCNCAGCTTTCNCNNC
ATNATNCTNCNNGAANTNTCANNCNTCTCCNCTTNAATGNNTTCCCATGNATTAA

Marked-Up Copy

NTNCCTCGNNNANAGCACTATCGTNNNNGAGNNNATTATNGNCNNTTTACNTCATG
TGGTCCANTNNCGTTNGNCGCNNNNAATNTTCGTNNNNNCNN

~~Sequence ID 1118~~ SEQ ID NO: 390

GGATTTTAGAGGAAGGCGCTNGGTTACATTGGAGAACTGGAGTGGTCTGGAGTTCC
ACGGTGTAGTGGACCAGAGGCCACCTCTCCTGGGCTTCTCAGTGTCTCGCCGGCGG
GGTTCGGCCTGAGCTGGATTGACATAGCCCTTGGCGGATTTAAACAACCTAAACAT
TAAGCAGTACAGCTGCCTCAAACCTTTGGGATTTTTCAGAATGACTGACACTGCCGA
AGCTGTTCCAAAGTTTGAAGAGATGTTTGCTAGTAGATTTCACAGAAAATGACAAGG
AGTATCAGGAATACCTGAAACGCCCTCCTGAGTCTCCTCCAATTGTTGAGGAATGG
AATAGCANAGCTGGTGGGAACCAAAGAAACAGAGGCAATCGGTTGCAAGACAACAG
ACAGTTCAGAGGCAGGGACAACAGATGGGGGTGGCCAAGTGACAATCGATCCAATC
AGTGGCATGGACGATCCTGGGGTAACAACACTACCCGCAACACAGACAAGAACCTTAC
TATCCCCAGCAATATGGACATTATGGTTACAACCAGCGGCCTCCTTACGGTTACTA
CTGATAGAAATGTTGGCAGCTTTTAGTAAAAGCATTTACTCTGTTACCATGAGAAA

~~Sequence ID 1125~~ SEQ ID NO: 391

NGACTGGCTCCCGAAAAGAAGGGTGGCGAGAAANAAAAGGGCCGTTCTGCCATGGA
CGAAGTGGTAACCCGCGAATACACCATCAACATTNACAAGCGCATCCATGGAGTGG
GCTTCAAGAANCCTGCACCTCGGGCACTCAAAGAGATTCGGAAATTTGCCATGAAG
GAGATGGGAACTCCATATGTGCGCATTGACACCAGGCTCAACAAANCTGTCTGGGC
CAAAGGAATAAGGAATGTGCCATACCGAATCCGTGTGCGGCTGTCCANAAAACGTA
ATGAGGATGAAGATTCACCAAATAAGCTNTATACTTTGGTTACCTATGTACCTGTT
ACCACTTTCAAAAATCTACAGACAGTCAATGTGGATGANAACNAATCGCTGATCGT
CAGATCAAANAAANT

~~Sequence ID 1139~~ SEQ ID NO: 392

nt:

503

CAGCACTGCCAGTGGAGATGGGCGTCACTACTGCTACCCTCATTTACCTGCGCTG
TGGACACTGAGAACATCCGCCGTGTGTTCAACGACTGCCGTGACATCATTCAGCGC
ATGCACCTTCGTACGTACGAGCTGCTCTAAGAAGGGAACCCCCAAATTTAATTA
GCCTTAAGCACAATTAATTAAGTGAACGTAATTGTACAAGCAGTTAATCACCC
ACCATAGGGCATGATTAACAAAGCAACCTTCCCTTCCCCCGAGTGATTTTGCGAA
ACCCCCTTTTCCCTTCAGCTTGCTTAGATGTTCCAAATTTAGAAAGCTTAAGGCGG
CCTACAGAAAAAGGAAAAAGGCCACAAAAGTTCCCTCTCACTTTCAGTAAAAATA
AATAAACAGCAGCAGCAAACAAATAAATGAAATAAAAGAAACAAATGAAATAAA

TATTGTGTTGTGCAGCATTAAAAAAATCAAATAAAAATTAAATGTGAGCAAAG

~~Sequence ID - 1148~~ SEQ ID NO: 393

nt:

587

TGAAAAATAAAGTTTTTATGTATATTCTACATATGTATATGTTGGTAGAAAGCAAA
AACGCTAGGTAAAAATAAATGTAATACAATTTTAGCTATGAACCAAAAAACCATTT
GTGGTGTGGATGCAAGAAAGTCTGGATGGGTGCAGAGTTCTCCATGTTTCACTTCT
GACATTTGAAAATACGCAGTTTGCATTTGATACGTCAAATGTTATTTTTTAAGAAAA
CCAATAAAATCATTA AAAACCGAAAAGGCAGTTTTGCTTGTTTTTACCTTAGTTGGA
GTTATCTGCAATTGCCGTATTAGTGTTTTAAGGAACTTGTAAGTAAGCTCCTTAGT
CCCCTTTAGAGCTACGAAACATGTCAATTTTACTTTTCTCCAGCTTTTTTGGAATCT
TATCTAAATTACCATGTAGAGTTCTGCATAGCTTCAAATTCCTTAGCCAATGTGG
TCTGTAAGTGTCTATCGATGAATTTACCGTTAATTGCCGTAGTATACTGTCCTGT
ACCGGATGTGAAGAGGAGCAACTCTGCACAGTGCACTGGTTGCTCCCATGGTAGGA
ANGAATGGCTTATCAATGGTCGGATTT

~~Sequence ID - 1160~~ SEQ ID NO: 394

nt:

650

GGAGGATGGAGCAGTGAGCGGGTCTGGGCGGCTGCTGGCAGCGCCATGGAGACGGT
ACAGCTGAGGAACCCGCCGCGCCGGCAGCTGAAAAAGTTGGATGAAGATAGTTTAA
CCAAACAACCAGAAGAAGTATTTGATGTCTTAGAGAACTTGAGAGAAGGGTGAGTG
TAAAGAACTATAGGTAGGTCATTGGGTCCCAGTCTTTTTCTGCCCCAGAAGAAG
CAGAAGGATATGAACCTTTCAGCATTTGTTCTAGGTGGGGTGGAAGGTAAATTTACA
GCTTGTGATGTCCTTCTTCGCTTTACTCCAATCCCTATTATAGACAGATTTAGTGA
TTCCTGGTCTTTTTTAACACGAAGAATATCTATTGTTTTCTCTTTTGTAGGATCTGT
ATGATTTTATCTACTTAACAGATAGCACTAATTAGATTAAAATTCTATAAGAACT
TTTTAATTTGCTGTTTATAATTTCTGATTGGTATGCAATAACTGTTTCAATGAAAA
TCAATGTAATTTAGTATTTTAATATTTGCACCTTTGTGAAATATAGTAAATAAATT
AAGCACTATCACACCTTCACAGCTACTTAGGAGATCCACAATCCTGGGTTGGGAG
CCAGTGGATTTCTGAAACACAGATTTGTTAATG

~~Sequence ID - 1165~~ SEQ ID NO: 395

nt:

502

CTCAAGTGAATCCTGGCTTCTTGGAAGCGCTTGCCTAGACGAGACACAGTGCATAA
AAACAACTTTTGGGGGACAGGTATGTTTTCTTGACGCTGCGGTTGTAAGGTCTTGG
CAAGACAAGCAGTGTGGCCAGAATTTTGAACCTTCTGATGAATGTGTAATGCAAAGG

Marked-Up Copy

ACCTTGTACATTTTTTTGTTTCAAGGTCCTCAAAATGAGCACATGAAGAGGTTGCT
GTGAAACTTTAAGTGGCCCTACTGCGCAGAAGCATTCAGATGTCACTTGATGATCT
GTAAGGGAACCTTGCTGATTTGGGAATGTGCTTAGGGAACACACATTCCTTTTGACA
GGGTCTGTCACTGGGTGGGTGATGAATTATACAGATGACATGTGCTTTTTTTTCTT
TTTTCAACCTCAATGGTATTCCTACAGGAAATGGATAACCATTTTAACTGTATTTT
TTGCAGCCCGTACCTTCTTGGGAATACAATTGTCTAACTTTTTATTTTTTGGTCT

~~Sequence ID - 1172~~SEQ ID NO: 396

nt:

648

CCACAATAATAAGAGAAAAACAGGAGCAAAAGGATATACAAAACCACCAGAAAACA
AATAACAAAGTGACAGGAGTAAGTCCTTAACTGGCAATAATAACCATGAATCTAAA
TGGATTCCATTTCCCCTTAAAAGATAAAGACATGCTGAATGGATAAAAAGCTGTC
ACCCAGTTATATGCTGCCTACAACAACTCACTTCACCTGTAAACATACATATGGA
TGGAAGAGAAAGGCATGGGAAAAGATACTCTACTCAAATGAAAACAAAAACCAAAC
AAAGGTGGCTATTCTTATATGAGATAATACAGACATTAAATCAAAAACCTGGAAACA
AACACAAAGTCATTGTATAATGATGAATTCAATTATATCATGATGAATTCAATTAT
ATCCTCCTTCCTGATCAATTCAGAAAGGAGGATATAATCTTTTTAAATATATATAC
ACCCAACACCAGAGCATATAAATATGTAAAGGAAGATAAAGGGAGTCCTGTGATCA
AGAATAAATATAACAATTATAAATATTTTATCTAAAGTGATAGATAGACTGTAATA
CAATAATAGGGTGGTGACATTAACACCCCCCTCTCACATTGGACTGATCATCTAGAA
GGGAGAAAAAGCTTTTATGATTGGAAAAGCCAT

~~Sequence ID 1178~~SEQ ID NO: 397

ATTGTGTTGGCCACCCGGAATTCGCGGCCGCGTCGACCTACGCACACGAGAACAT
GCCTCTCGCAAAGGATCTCCTTCATCCCTCTCCAGAAGAGGAGAAGAGGAAACACA
AGAAGAAACGCCTGGTGCAGAGCCCCAATTCCTACTTCATGGATGTGAAATGCCCA
GGTGAGGAGACGGCTTGCTGTAGTGGGGAAAGCACTGGACCTCAACAGTTGGAAAA
TGTTGTAGTGTTAGCTGTCTCGTATCCTTGAAGCTGTGCAGCAGCTTCAGTTTCTT
CGCCTGTGGAATAATTTTCCCTGATACTCTTAAAATTTGAATGTATGAGACTGGC
AAAGTTTTGCATCTTAGGAGGAGTGATTCATTTACCGTGATCTCTCATCACATTT
CACATACAACCCCTACGTTTTTTTTGTGTTGGGAAACAATGTAATGGATGATGAGTT
GGGCATAAGTGCAGGAAAGACGGGTGTAATAGAGGAAAAAATGTTATCTGCTTTT
CTTTCAGGATGCTATAAAATCACACGGTCTTTAGCCATGCACAAACGGTAGTTTTT
GTGTGTTGGCTGCTCCACTGTCCTCTGCCAGCCTACAGGAGGAAAAGCAAGGCTTA
CAGAAGGATGTTCTTCAGGAGGAAGCAGCACTAAAAGCACTCTGAGTCAANATGA
GTGGGAAACCATCTCAATAAACACATTTTGGAT

~~Sequence ID - 1180~~ SEQ ID NO: 398

nt:

622

CTTTTCCTCCCGCTGTCCCCCACGGGAGGGGACTGCTCTCCCCCGCTGCATCCTTT
CTGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAACCT
CGACATTCGGGAAGTGTTTTTGAGAAGTCTCGGTTCGGTAAGGGAAGTCTTCCAAGT
CCGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTTCGGCCACCCCCCAGATGAG
GCAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTTCTCTCAGCT
TCCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGTCCATGCTGCCA
TCAAGGCATTTATTGCAGTGTACTATTTGCTTCCAAAGGATCAGGCCCTGAGAACA
ATGACCTTATTTCTACAACAGTGTCTGGGTTGCGTGCCAGCAGATGCCTCAGATA
CCAAGAGATAACAAAGCTGCAGCTCTTTTGATGCTGACCAAGAATGTGGATTTTGT
GAAGGATGCACATGAAGAAATGGAGCAGGCTGTGGAAGAATGTGACCCTTACTCTG
GCCTCTTGAATGATACTGAGGAGAACAACCTCTGACANCCACAATCATGAGGATGAT
GTGTTG

~~Sequence ID - 1181~~ SEQ ID NO: 399

nt:

155

CGCCACTTATCCAGTGAACCACTATCACGAAAAAACTCTACCTCTCTATACTAAT
CTCCCTACAAATCTCCTTAATTATAACATTACAGCCACAGAACTAATCATATTAA
AA

~~Sequence ID - 1182~~ SEQ ID NO: 400

CATTGTGTTGGCNCCCGGAATTCGCGGCCGCGTCGACTTTTTGTGTTGTTTGGAG
CAGAAATACTAAAGAAGATTCCGGGCCGAGTATCCACAGAAGTGACGCAAGGCTC
TCCTTTGATAAAGATGCGATGGTGGCCAGAGCCAGGCGGCTCATCGAGCTCTACAA
GGAAGCTGGGATCAGCAAGGACCGAATTCTTATAAAGCTGTCATCAACCTGGGAAG
GAATTCAGGCTGGAAAGGAGCTCGAGGAGCAGCACGGCATCCACTGCAACATGACG
TACTCTTCTCCTTCGCCCAGGCTGTGGCCTGTGCCGAGGCGGGTGTGACCCTCAT
CTCCCCATTTGTTGGGCGCATCCTTGATTGGCATGTGGCAAACACCGACAAGAAAT
CCTATGAGCCCCCTGGAAGACCCTGGGGTAAAGAGTGTCACTAAAATCTACAACCTAC
TACAAGAAGTTTAGCTACAAAACCATTTGTCATGGGCGCCTCCTTCCGCAACACGGG
CGAGATCAAAGCACTGGCCGGCTGTGACTTCCTCACCATCTCACCCAAGCTCCTGG
GAGAGCTGCTGCAGGACAACGCCAAGCTGGTGCCTGTGCTCTCAGCCAAGGCGGCC
CAAGCCAGTGACCTGGAAAAAATCCACCTGGATGAGAAGTCTTTCCGTTGGTTGCA
CAACGAGGACCAGATGGCTGTGGAGAAG

~~Sequence ID - 1183~~ SEQ ID NO: 401

nt:

479

CGTGGCAGCCATCTCCTTCTCGGCATCATGGCCGCCCTCAGACCCCTTGTGAAGCC
CAAGATCGTCAAAAAGAGAACCAAGAAGTTCATCCGGCACCAGTCAGACCGATATG
TCAAAATTAAGCGTAACTGGCGGAAACCCAGAGGCATTGACAACAGGGTTTCGTAGA
AGATTCAAGGGCCAGATCTTGATGCCCAACATTGGTTATGGAAGCAACAAAAAAC
AAAGCACATGCTGCCCAGTGGCTTCCGGAAGTTCCTGGTCCACAACGTCAAGGAGC
TGGAAGTGCTGCTGATGTGCAACAAATCTTACTGTGCCGAGATCGCTCACAATGTT
TCCTCCAAGAACCGCAAAGCCATCGTGGAAAGAGCTGCCCAACTGGCCATCAGAGT
CACCAACCCCAATGCCAGGCTGCGCAGTGAAGAAAATGAGTAGGCAGCTCATGTGC
ACGTTTTCTGTTTAAATAAATGTAAAACTG

~~Sequence ID - 1185~~ SEQ ID NO: 402

nt:

628

CTTTGATTACCTTTGAGTATTAGGTGAAAGCTTCTCTGTGCTTGATTGAACATTG
TGATGATGTTGATTGGGTCATGTCAGATTTAGACAGTGTTGTGTTTAAGATAAATG
TTTAATGGCTCTTAGCAGTGTTTCATGCCTCCCCTTTTCCCCTGATACTTTAAAAAC
AGAATATACAGAAAAGGGGAGTTGGGTGAAGAATCACCATATTCTCATTACCAGAG
TAGTGTCTACCAGCTGTTTTACATTTTTCTGTTTCCTTCTGTCTTGGAATCCTT
TTTTTAGATCCTTGTAATACTAGTAAAGATATTCCACTCTGTGTTGTAAGCATTTT
TCCATTTTGCTCCATGGTCTTCATAATGCCCTGTGGTCCTTTATTAAGGGGATGCA
CCATGTAGAGGTGAAAGGCTTTCCTTGACTTGGCCACCATTTCTGTATTTTCCTTA
GAGGAGGAGGTTTCCAACATTTCTTTTTTAGAGACAGAGTCTCGTTCTGACACGCA
GGCAGGAGTGCAGTGGCATGATAACAGCTCACTGCAGCCTCGAACTCCTGGGCTCA
AGTTATCCTCCCACCTCAGCTTCCTGAGTAGCTAGGACTGCAGGTGCCTGCCACCA
CACCCAGCTAAT

~~Sequence ID - 1186~~ SEQ ID NO: 403

nt:

494

CAGCCCTCCGTACCTCTTCACCGCACCCCTCGGACTGCCCCAAGGCCCCCGCCGCC
GCCTCCAGCGCCGCGCAGCCACCGCCGCCGCGCCGCTCTCCTTAGTCGCCGCCA
TGACGACCGCGTCCACCTCGCAGGTGCGCCAGAACTACCACCAGGACTCAGAGGCC
GCCATCAACCGCCAGATCAACCTGGAGCTCTACGCCTCCTACGTTTACCTGTCCAT
GTCTTACTACTTTGACCGCGATGATGTGGCTTTGAAGAACTTTGCCAAATACTTTT
TTCACCAATCTCATGAGGAGAGGGGAACATGCTGAGAACTGATGAAGCTGCAGAA

Marked-Up Copy

CCAACGAGGGTGGCCGAATCTTCCTTCAGGATATCAAGAAACCAGACTGTGATGAC
TGGGAGAGCGGGCTGAATGCAATGGAGTGTGCATTACATTTGGAAAAAATGTGAA
TCAGTCACTACTGGAAGTGCACAACTGGCCACTGACAAAAATGAC

~~Sequence ID - 1188~~ SEQ ID NO: 404

nt:

599

GGGAGACAAGCCCAGCCTTTCGGCGAGNATACGTCTAACCCTGTGCAACAGCCACT
ACATTACTTCAAAGTGAAGATCCTTCCTTTTGAGGGAGCAAGTCCTTCCCTTTCATT
TTTTCCAGTCTTCCTCCCTGTGTATTCAATTCTCATGATTATTATTTAGTGGGGGCG
GGGGTGGGAAAGATTACTTTTTCTTTATGTGTTTGACGGGAAACAAAAGTAAAGTAA
AATCTACAGTACACCACAAGGGTCACAATACTGTTGTGCGCACATCGCGGTAGGGC
GTGGAAAGGGGCGAGCCANAGCTACCCGCAGAGTTCTCAGAATCATGCTGAGAGAG
CTGGAGGCACCCATGCCATCTCAACCTCTTCCCCGCCCCGTTTTACAAAGGGGGAGG
CTAAAGCCCAGAGACAGCTTGATCAAAGGCACACAGCAAGTCAGGGTTGGAGCAGT
AGCTGGAGGGACCTTGTCTCCCAGCTCAGGGCTCTTTCCTCCACACCATTTCAGGTC
TTTCTTTCCGAGGCCCCCTGTCTCAGGGTGAGGTGCTTGAGTCTCCAACGGCAAGGG
AACAAGTACTTCTTGATACCTGGGATACTGTGCCAGAG

~~Sequence ID - 1189~~ SEQ ID NO: 405

GGGAGACAAGCCCAGCCTTTCGGCGAGATACGTCTAACCCTGTGCAACAGCCACTA
CATTACTTCAAAGTGAAGATCCTTCCTTTTGAGGGAGCAAGTCCTTCCCTTTCATTT
TTTCCAGTCTTCCTCCCTGTGTATTCAATTCTCATGATTATTATTTAGTGGGGGCG
GGGTGGGAAAGATTACTTTTTCTTTATGTGTTTGACGGGAAACAAAAGTAAAGTAA
ATCTACAGTACACCACAAGGGTCACAATACTGTTGTGCGCACATCGCGGTAGGGCG
TGGAAAGGGGCGAGCCAGAGCTACCCGCAGAGTTCTCAGAATCATGCTGAGAGAGC
TGGAGGCACCCATGCCATCTCAACCTCTTCCCCGCCCCGTTTTACAAAGGGGGAGGC
TAAAGCCCAGAGACAGCTTGATCAAAGGCACACAGCAAGTCAGGGTTGGAGCAGTA
GCTGGAGGGACCTTGTCTCCCAGCTCAGGGCTCTTTCCTCCACACCATTTCAGGTCT
TTCTTTCCGAGGCCCCCTGTCTCAGGGTGAGGTGCTTGAGTCTCCAACGGCAAGGGA
ACAAGTACTTCTTGATACCTGGGATACTGTGCCAGAGCCTCGAGGAGGT

~~Sequence ID - 1190~~ SEQ ID NO: 406

GTTTAAATTTGACAAACTAAAGCTNATNACTGCTATAAGAGTAATAACTGCTCATT
TTCCATAACTCATTCTTAAAGTTTTAGTAATGTAAAAGTTATTTTTTTGCAGTAAG
TTATAATGATAGAAGCTTACATGTTTTTTCATGCCTCATCTGTTTCCCTTAAAC
TATAATTATCAGTAAAGTCCTGTGGTATTTTTCAATTTGTAAGAACTAGGCTATA

Marked-Up Copy

TATACATTGGGAAAAACAGCCTTCATTTGTCAATGCACTAGTGTTCCAAAGGTTTC
TGGTAATTGTGTGCTATTGCTTTTTTGTGACTTGCAAAAAAAAAAAAAAAAAAATT
ACTATGACTTGNGGTAGCCCTGCAACCTTCGGAAGTGCTTAGCCCAGTCTGACCAT
ACATTTATATTTANAATGCTTAGGTAAATAAATAATATGCCTAAACCCAATGCTAT
AAGATACTATATAATATCTCATAATTTTAAAAATCACTGTTTTGTATAATAATAAA
ACAAGGCAGGCAAGCTGTTCTACAATGACTGTTGGTAAGGGTGCTGAGGAAGAAAA
ACAAACAATCTTGATTCAGGGATAGTGAATAGACAAAAAATGTCCTAATCAATGAA
GCTGTGTGATGATTCTGATTGACAGAGA

~~Sequence ID 1191~~SEQ ID NO: 407

GTGCAAAGTGTTATATCCACTTTCAACAAAGAGAGAAGCTGAAAAGCTAACCCAAT
GTTAATTTTGGATCACACACATTCAGTGTAGACTTTAAGATTTTACTTCTGTTGGA
GTAGCTATATTATTTCTAGTTAAAAAAGTCTCTATATACATATTTATTTGTTTTTC
TACTTGTTTAATATTTTTCTCTTCCAATTAGGAACTCAATATGGAATAAAAAATAT
TTAAATGTATTTTACTCAAACGTGTGTGTATATATGTTTGTGTGCATGATAAGGAG
AGTGAGAGCAAGAGTAAGAGAGAGAGAGCACGCATAGATGGAAGCACACATTTAAT
GTCTATGAAATGAGAAAACATTAAGGCTAAGATATTTTTCTTCTGAACTAGCAGA
TTGTATCAATGGCTGGTCACTTAAATTAATCAGTTTGTAAGATATTTAAAAGGTA
TGTCTACCTTCTTGCAATTAATTTGATTATGTTCTAATGGCATGGCAAGAGAAATG
AAAGAAGATAACTAAAAGTTAAAAGTCGTTGCATGTTTTTGTGTCAGCATACCTTT
CTTTCAGGCTACCGAATAACCTTGATTGACATTGGATTAGTAGTAGAATACCTCAT
TGGTAGAGCATATCGCAGCANCTACACTAGAAAACAT

~~Sequence ID 1192~~SEQ ID NO: 408

GTCTGGAACCTCCAGACCTCAGGTGATACCCCTGCCTCAGCCTCCCAATGTGCTGGG
ATTACAGCTGTGAAGCCACCGCGCCCGGCTGCTGTGATAGTTGAGATGTAAACCAA
AAATAAAATTCTAAGCCACCCAATCCGACTGAATGGACCCTTCCTGTTGAGCAAGG
ACATTCCAAAGTAACTGAAAAGACCAGCTTAGGCCATGATGGGAAGGGGAGGTGT
CAACATGCCTCATTCTACCTTCCCTCTGGAATCCAGACACAAGTACCAGCAT
TAACATTAAAACAGAGATCTTAAGCTGGGCACGGTGGCTCATGCCTGTAATCCCAG
CACTTTGGGAGGCCAAGGTGGGATCACCTGAGGTCGGAAGTTCAAGACCAGCCTGG
CCGGTATGGTGAAGCCATGTCTCTACTGAAAATGCAAAATTGGCCGGACATTGTGG
TGCA

~~Sequence ID 1193~~SEQ ID NO: 409

TNCNTTTTTTTTCCCNCGGGAAAGCGCGCCATTGTGTTGGTCCCCGGGAATTCGCG

Marked-Up Copy

GCCGCGTCGACGAGAAATGGCTTGAACCCAGTAGGCAGAGGTTGTAGTGAGCCCAG
AATNGGNCACCTGCACNTTTANCCNTGGGTGACAAAANTGAAAACCTTTGTCTNAAA
AAAAAAAAAAAAAAAAATTTAANTNAAATNAAAAANCCTTTNCNTTNTTTTTTNAAAN
NGGGGGGGGNNTTTTTNGGGNTTNGNNNTGGTAAAAANTNNNTTTTTTTTTTTTA
GGGGCCNANNCCCCNTTTTANAAAAANCCNGNTTTTNAAAAAANTTTTTTNCCNCN
NTTNGGGGGGGGGGNTTTTNANCNTNTTNGGGGGGGGNCCCCCTNTTANNACCNNC
AAANTTTTTANTTTTTTGNNAANNCCCCCTTTTTTNNTTTTTTTTTTGNNGGGGGG
GGGNNGCCCCCNCCCTTTNGGGGGGGGGGNTTNGNAAAANNACTTTTNAAAANNA
AGGGNNGGGGGNANATNNCCCCCCCCNGGNTTTTTTTTTTTAAAAANTNAANNGGGG
GGGNNTCTNANTNGGGGCNCCCANNGGGGNTTANAANNATTTTCTNCCCAAACCC
CCNGNTTTTATNNCCCCCCCCCCCCCNCNNNNGAANGGGNGGNCNTTTTTTTTTTAT
TTTNGGNGGGNAAAAAANTTTNAAAAANNANATNTTTTTTCCCCCCCCCCCCCNC
TTTTNGGNAAANCCNNGGGGGGNTCCTTTTTNAAANNNNCCCCCAAAAAAANTTT
TTTTNTNTNTTTTTTCTCTNGGGGNCNNANTTNTANANTTTTNCNCNAAAAAAA
ANGGGNCCCCTTTTTTTNCNGGNNGGNNCCCAAAANNTTTTTTTTNAAAAAAAA
AAAA

~~Sequence ID 1195~~ SEQ ID NO: 410

GTTCGTGACNTTCGGAGCTACCTGACAGAGCAGAGTCAACCAGGNTCTGCCCAAAG
AGAGTGTTAGGCCTGAGCTTGAGAGCCCTGGAGAGACGTGTGCACAAAATGTGACC
TGAGGCCCTAGTCTAGCAAGAGGACATAGCACCCCTCATCTGGGAATAGGGAAGGCA
CCTTGCAGAAAATATGAGCAATTTGATATTAACATAACATCTTCAATGTGCCATAGA
CCTTCCCACAAAGACTGTCCAATAATAAGAGATGCTTATCTATTTTA

~~Sequence ID 1196~~ SEQ ID NO: 411

nt:

412

GTCGACGCGGCCGCGGTCGCTGGAGNCGATCAACTCTAGGCTCCAACCTCGTTATGA
AAAGTGGAAGTACGTCCTGGGGTACAAGCAGACTCTGAAGATGATCAGACAAGGC
AAAGCGAAATTGGTCATTCTCGCTAACAACCTGCCAGCTTTGAGGAAATCTGAAAT
AGAGTACTATGCTATGTTGGCTAAAACCTGGTGTCCATCACTACAGTGGCAATAATA
TTGAACTGGGCACAGCATGCGGAAAATACTACAGAGTGTGCACACTGGCTATCATT
GATCCAGGTGACTCTGACATCATTAGAAGCATGCCAGAACAGACTGGTGAAAAGTA
AACCTTTTCACCTACAAAATTTACCTGCAAACCTTAAACCTGCAAAATTTTCCTT
TAATAAAATTTGCTTGTTTT

~~Sequence ID 1197~~ SEQ ID NO: 412

Marked-Up Copy

CCGCCAACATGGGCCGCGTTTCGCACCAAAACCGTGAAGAAGGCGGCCCGGGTCATC
ATAGAAAAGTACTACACGCGCCTGGGCAACGACTTCCACACGAACAAGCGCGTGTG
CGAGGAGATCGCCATTATCCCCAGCAAAAAGCTCCGCAACAAGATAGCAGGTTATG
TCACGCATCTGATGAAGCGAATTCAGAGAGGCCAGTAAGAGGTATCTCCATCAAG
CTGCAGGAGGAGGAGAGAGAAAGGAGAGACAATTATGTTCTGAGGTCTCAGCCTT
GGATCAGGAGATTATTGAAGTAGATCCTGACACTAAGGAAATGCTGAAGCTTTTGG
ACTTCGGCAGTCTGTCCAACCTTCAGGTCCTCAGCCTACAGTTGGGATGAATTTTC
AAAACGCCTCGGGGACCTGTTTGAATTTTTTCTGTAGTGCTGTATTATTTTCAATA
AATCTGGGACAA

~~Sequence ID 1198~~ SEQ ID NO: 413

CAGAGGTGGGAGGATTGCTTCAGTTCAAGAGTTTGAGACCAGCCTGGGTAACATGG
CGAAACCCTGTCTTTACAAAAATGCAAACCTTTGCCGCATGTGTTGGGGTGCGCC
TGTAGTCCCAGCTTCTCGGGAGGCTGAGGTGGGGGGACCACCTGAGCCATGGAGGT
TGAGGCTGCAGTGAGCCGTGATACCACCACTGTACTCTAGCCTGGGCCATAGAGTG
AGACACCCTGCCTCAGAAATA

~~Sequence ID 1199~~ SEQ ID NO: 414

nt:

439

CCCATCCCCTCGACCGCTCGCGTCGCATTTGGCCGCCTCCCTACCGCTCCAAGCCC
AGCCCTCAGCCATGGCATGCCCCCTGGATCAGGCCATTGGCCTCCTCGTGGCCATC
TTCCACAAGTACTCCGGCAGGGAGGGTGACAAGCACACCCTGAGCAAGAAGGAGCT
GAAGGAGCTGATCCAGAAGGAGCTCACCATTGGCTCGAAGCTGCAGGATGCTGAAA
TTGCAAGGCTGATGGAAGACTTGGACCGGAACAAGGACCAGGAGGTGAACTTCCAG
GAGTATGTCACCTTCCTGGGGGCCTTGGCTTTGATCTACAATGAAGCCCTCAAGGG
CTGAAAATAAATAGGGAAGATGGAGACACCCTCTGGGGGTCTCTCTGAGTCAAAT
CCAGTGGTGGGTAATTGTACAATAAATTTTTTTTTTGGTCAAATTTAA

~~Sequence ID 1200~~ SEQ ID NO: 415

nt:

526

CTGGAGACGACGTGCAGAAATGGCACCTCGAAAGGGGAAGGAAAAGAAGGAAGAAC
AGGTCATCAGCCTCGGACCTCAGGTGGCTGAAGGAGAGAATGTATTTGGTGTCTGC
CATATCTTTGCATCCTTCAATGACACTTTTGTCCATGTCACTGATCTTTCTGGCAA
GGAAACCATCTGCCGTGTGACTGGTGGGATGAAGGTAAAGGCAGACCGAGATGAAT
CCTCACCATATGCTGCTATGTTGGCTGCCCAGGATGTGGCCAGAGGTGCAAGGAG
CTGGGTATCACCGCCCTACACATCAAACCTCCGGGCCACAGGAGGAAATAGGACCAA

Marked-Up Copy

GACCCCTGGACCTGGGGCCCAGTCGGCCCTCANAGCCCTTGCCCGCTCGGGTATGA
AGATCGGGCGGATTGAGGATGTCACCCCCATCCCCTCTGACAGCACTCGCAGGAAG
GGGGGTCGCCGTGGTCGCCGTCTGTGAACAAGATTCCCTCAAAATATTTTCTGTAA
TAAATTGCCTTCATGTAACTG

~~Sequence ID — 1201~~ SEQ ID NO: 416

nt:

613

CTTAAGTATGCCCTGACAGGAGNATGAAGTAAAGAAGATTTGCATGCAGCGGTTCA
TTAAAATCGATGGCAAGGTCCGAACCTGATATAACCTACCCTGCTGGATTCATGGAT
GTCATCAGCATTGACAAGACGGGAGAGAATTTCCGTCTGATCTATGACACCAAGGG
TCGCTTTGCTGTACATCGTATTACACCTGAGGAGGCCAAGTACAAGTTGTGCAAAG
TGAGAAAGATCTTTGTGGGCACAAAAGGAATCCCTCATCTGGTGACTCATGATGCC
CGCACCATCCGCTACCCCGATCCCCTCATCAAGGTGAATGATACCATTGAGATTGA
TTTAGAGACTGGCAAGATTACTGATTTTCATCAAGTTCGACACTGGTAACCTGTGTA
TGGTGACTGGAGGTGCTAACCTAGGAAGAATTGGTGTGATCACCAACAGAGAGAGG
CACCTGGATCTTTTGACGTGGTTCACGTGAAAGATGCCAATGGCAACAGCTTTGC
CACTCGACTTTCCAACATTTTTGTTATTGGCAAGGGCAACAAACCATGGATTTCTC
TTCCCCGAGGAAAGGGTATCCGCCTCACCATTGCTGAAGAGAGAGACAAAAGA

~~Sequence ID — 1202~~ SEQ ID NO: 417

GGAATTTCGCGCCGCGTCGACCTCTGCTCGAATTGACAGAAAAGGATTCTGTGAAG
AGTGATGAGATTTCCATCCATGCTGACTTTGAGAATACATGTTCCCGAATTGTGGT
CCCCAAAGCTGCCATTGTGGCCCGCCACACTTACCTTGCCAATGGCCAGACCAAGG
TGCTGACTCAGAAGTTGTCATCAGTCAGAGGCAATCATATTATCTCAGGGACATGC
GCATCATGGCGTGGCAAGAGCCTTCGGGTTGAGAAGATCAGGCCTTCTATCCTGGG
CTGCAACATCCTTCGAGTTGAATATTCCTTACTGATCTATGTTAGCGTTCCTGGAT
CCAAGAAGGTCATCCTTGACCTGCCCCTGGTAATTGGCAGCAGATCAGGTCTAAGC
AGCAGAACATCCAGCATGGCCAGCCGAACCAGCTCTGAGATGAGTTGGGTAGATCT
GAACATCCCTGATACCCCAGAAGCTCCTCCCTGCTATATGGATGTCATTCTGAAG
ATCACCGATTGGAGAGCCCAACCACTCCTCTGCTAGATGACATGGATGGCTCTCAA
GACAGCCCTATCTTTATGTATGCCCTGAGTTCAAGTTCATGCCACCACCGACTTA
TACTGAGGTGGATCCCTGCATCCTCAACAACAATGTGCAGTGAGCAT

~~Sequence ID — 1203~~ SEQ ID NO: 418

nt:

692

TGCAGAGGGGTCCATACGGCGTTGTTCTGGATTCCCGTCGTAACTTAAAGGGAAAC

Marked-Up Copy

TTTCACAATGTCCGGAGCCCTTGATGTCCTGCAAATGAAGGAGGAGGATGTCCTTA
AGTTCCTTGCAGCAGGAACCCACTTAGGTGGCACCAATCTTGACTTCCAGATGGAA
CAGTACATCTATAAAAGGAAAAGTGATGGCATCTATATCATAAATCTCAAGAGGAC
CTGGGAGAAGCTTCTGCTGGCAGCTCGTGCAATTGTTGCCATTGAAAACCCTGCTG
ATGTCAGTGTTATATCCTCCAGGAATACTGGCCAGAGGGCTGTGCTGAAGTTTGCT
GCTGCCACTGGAGCCACTCCAATTGCTGGCCGCTTCACTCCTGGAACCTTCACTAA
CCAGATCCAGGCAGCCTTCCGGGAGCCACGGCTTCTTGTGGTACTGACCCCAGGG
CTGACCACCAGCCTCTCACGGAGGCATCTTATGTTAACCTACCTACCATTGCGCTG
TGTAACACAGATTCTCCTCTGCGCTATGTGGACATTGCCATCCCATGCAACAACAA
GGGAGCTCACTCAGTGGGTTTAATGTGGTGGATGCTGGCTCGGGAAGTTCTGCGCA
TGCGTGGCACCATTTCCTCGTGAACACCCATGGGAGGTCATGCCTGATCTGTACTTC
TACAGAGATCCTGAAGAGAT

~~Sequence ID 1204~~SEQ ID NO: 419

TTTTTTTTTTTTTCTGCGGGAAAGCGCGCCATTGTGTTGGTACCCGGGAAATTCTG
CGGCCGCGTCGACACAGGCCCCAGCATCAAGATCTGGGATTTAGAGAGGAAAGATC
ATTGTAGATGAACTGAAGCAAGAAGTTATCAGTACCAGCAGCAAGGCAGAACCACC
CCAGTGCACCTCCCTGGCCTGGTCTGCTGATGACACAGGTTGGGCNNGNNCNCNNG
GGNGGNNNNGNNNNGCNGNNGGNNCNGNNNNCNCNNNNNGCNCNNNGNNNTNNNCNN
GNNCNCNNNNNNNNNNNNNNNNNGNTCNCNNGNCGNGGGGCCNNGNCGNCGCGGNCGCG
NNTNNNNGGGTNCNNNCNCNNNGGCGCGC

~~Sequence ID 1205~~SEQ ID NO: 420

CAGACTCTGACCCAGCCTCAGTCCTAACTCCTGGGGCTGGGCTGAGGGGAACAAGC
ATTTGCTGAACTTGAAAAAACAAAGCAAATCAAAAACAGGAAAAAATTGTACCTG
GTACTTTTTTTTAGAAAAAAGATTAAAAAAGAAAGATAAATTCTTGTTTGGAAA
CTTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA
AAAAAATTTTAACTC
TNNNNNTNNCNCNANTAANNCANNTCNANNNNANNNNAATTACTTNNANGTNNNTC
ACN

~~Sequence ID 1207~~SEQ ID NO: 421

nt:

642

ACGAGAAGCCAGATACTAAAGAGAAGAANCCCGAAGCCAAGAAGGTTGATGCTGGT
GGCAAGGTGAAAAAGGGTAACCTCAAAGCTAAAAAGCCCAAGAAGGGGAAGCCCCA
TTGCAGCCGCAACCCTGTCCTTGTGAGAGGAATTGGCAGGTATTCCCGATCTGCCA

Marked-Up Copy

TGTATTCCANAAAGGCCATGTACAAGAGGAAGTACTCAGCCGCTAAATCCAAGGTT
GAAAAGAAAAAGAAGGAGAAGGTTCTCGCAACTGTTACAAAACCAGTTGGTGGTGA
CAAGAACGGCGGTACCCGGGTGGTTAACTTCGCAAAATGCCTAGATATTATCCTA
CTGAAGATGTGCCTCGAAAGCTGTTGAGCCACGGCAAAAAACCCTTCAGTCAGCAC
GTGAGAAAACCTGCGAGCCAGCATTACCCCCGGGACCATTCTGATCATCCTCACTGG
ACGCCACAGGGGCAAGAGGGTGGTTTTCTGAAGCAGCTGGCTAGTGGCTTATTAC
TTGTGACTGGACCTCTGGTCCTCAATCGAGTTCCTCTACGAAGAACACACCAGAAA
TTTGTCATTGCCACTTCAACCAAAATCGATATCAGCAATGTAAAAATCCCAAAACA
TCTTACTGATGCTTACTTCAAAAAGA

~~Sequence ID 1208~~ SEQ ID NO: 422

CCCTATACCTTCTGCATAATGAATTANCTAGAAATAACTTTGCAAGGGAGAGCCAA
AGCTAAGACCCCCGAAACCAGACGAGCTACCTAAGAACAGCTAAAAGAGCACACCC
GTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGGCGACAAACCTACCGAG
CCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAACTTTAAATTTGCCCA
CAGAACCCTCTAAATCCCCTTGTAATTTAACTGTTAGTCCAAAGAGGAACAGCTC
TTTGACACTAGGAAAAAACCTTGTAAGAGAGAGTAAAAAATTTAACACCCATAGTA
GGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACACCCACTACCTAA
AAAATCCCAAACATATAACTGAACTCCTCACACCCAATTGGACCAATCTATCACCC
TATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCCTCCGCATAAG

~~Sequence ID 1209~~ SEQ ID NO: 423

nt:

620

CTCTCCTGTCAACAGCGGCCAGCCTCCCAACTACGAGAATGCTCAAGGAGGAGCAG
GAAGTGGCTATGCTGGGGGCGCCCCACAACCCTGCTCCCCCGACGTCCACCGTGAT
CCACATCCGCAGCGAGACCTCCGTGCCCCGACCATGTTCGTCTGGTCCCTGTTCAACA
CCCTCTTCATGAACACCTGCTGCCTGGGCTTCATAGCATTCGCCTACTCCGTGAAG
TCTAGGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGGCCTATGCCTCCAC
CGCCAAGTGCCTGAACATCTGGGCCCTGATTTTGGGCATCTTCATGACCATTCTGC
TCGTCATCATCCCAGTGTTGGTTCGTCCAGGCCAGCGATAGATCAGGAGGCATCAT
TGAGGCCAGGAGCTCTGCCCCGTGACCTGTATCCCACGTACTCTATCTTCCATTCCCT
CGCCCTGCCCCCAGAGGCCAGGAGCTCTGCCCTTGACCTGTATTCCACTTACTCCA
CCTTCCATTCCCTCGCCCTGTCCCCACAGCCGAGTCCTGCATCAGCCCTTTATCCTC
ACACGCTTTTCTACAATGGCATTCAATAAAGTGTATATGTTTCTGGTGCTGCTGTG
ACTT

Marked-Up Copy

~~Sequence ID 1210~~ SEQ ID NO: 424

TTCGTAATTAGAATACTGTTTGGACTTGCTCAACAAGCACCTTATCTTAACAAAA
GTAAGTTATAGAAAAGGGAGACATTCATTTAACTTCAAGCCCATATTATTCTTAAA
AGCTGACTCTTGAAATAGTATTTATTGAGTCATAGTGGAGTCATGGGACTTTTTTAA
GGGCCGGAAGGGACTATTTAGATCATCCAGTCCCACCCTGTCATTTTATGGAGGAG
GAAACTGAGGCCTAGATAAGATAACCAGTTAGTGGGTCCACTGACCTTTAGGACAG
TAGTCTATCCGTAAGAGACAACATGGAGAAAGAAATACAACGTTTTTATAGTGAAT
TATCATCTTACAAAGAATATTCTTCCCATATCGCACTTTTAAAAAGTGGGTACCTT
AGTCAAATAGGAGAAAAAACCACTTGAGTAGTTTCATCCTCAGGTTTTAGGTGAGG
AAACTGATACTCAGATTAAATAACTTTAAGCACACAGAGCCTGAATGATAGTCTTA
TTTGAGCTCATCTGTGCTTTTAATGTGTACTACGTTAGGTGTTTTCACTTGCATTT
CCTTTAGTCTTATTTGAGCTCATCTGTGCTTTTAATGTGTACTACGTTAGGTGTTT
TCACTTGCATTTCTTGTGTTGACGTTGACAATAAATCGTGAAGCTGCCTTATCTAA
GGAAGTCCTAAAGTAAATCATTGGAACACA

~~Sequence ID 1211~~ SEQ ID NO: 425

CCATTGTGTTGGNACCCGGAATTTCGCGGCCGCGTCGACGGAGTTTTACCTTATTA
CACTTTAATCTCTGGATTTACCCCATCTCATTTCTCTTTTAGGAAAAGTGTGTA
TGTGGTGGCTTTGATGGTTCTCATGCCATCAGTTGTGTGGAAATGTATGATCCAAC
TAGAAATGAATGGAAGATGATGGGAAATATGACTTCACCAAGGAGCAATGCTGGGA
TTGCAACTGTAGGGAACACCATTTATGCAGTGGGAGGATTTCGATGGCAATGAATTT
CTGAATACGGTGGAAAGTCTATAACCTTGAGTCAAATGAATGGAGCCCCTATACAAA
GATTTTCCAGTTTTAACAAATTTAAGACCCTCTCAAACCTAACAGGCTTAGTGATGT
AATTATGGTTAGCAGAGGTACACTTGTGAATAAAGAGGGTGGGTGGGTATAGATGT
TGCTAACAGCAACACAAAGCTTTTGCATATTGCATACTATTAAACATGCTGTACAT
ACTTTTTGGGTTTATTTGGAAAGGAATGCAAAGATGAAGGTCTGTTTTGTGTACTT
TTAAGACTTTGGTTATTTTACTTTTTGGAAAAGAATAAACCAAGAATTGATTGGGC
ACATCATTTCAAGAAG

~~Sequence ID 1212~~ SEQ ID NO: 426

nt:

374

AGAGCAGCAGCCATGGCCCTACGCTACCCTATGGCCGTGGGCCTCAACAAGGGCCA
CAAAGTGACCAAGAACGTGAGCAAGCCCAGGCACAGCCGACGCCGCGGGCGTCTGA
CCAAACACACCAAGTTCGTGCGGGACATGATTCGGGAGGTGTGTGGCTTTGCCCCG
TACGAGCGGCGGCCATGGAGTTACTGAAGGTCTCCAAGGACAAACGGGCCCTCAA

Marked-Up Copy

ATTTATCAAGAAAAGGGTGGGGACGCACATCCGCGCCAAGAGGAAGCGGGAGGAGC
TGAGCAACGTACTGGCCGCCATGAGGAAAGCTGCTGCCAAGAAAGACTGAGCCCCT
CCCCTGCCCTCTCCCTGAAATAAAGAACAGCTTGACAG

~~Sequence ID - 1213~~ SEQ ID NO: 427

nt:

567

GAATTATTGACTTTGAATTGCATTTTCAGTACCATGAAGTCAAAGTCAGTGGTGTAT
TTGCTCATTTGTTCATTCTTTCTTTCCACCAACATTACTGCCTGCAGAGCCAGAG
GTGAGTGCAGAAATCCTGTCAATTCGTCACTTGTGGACAACCTGCAGCTTGCCACA
GCCTACAGTTCCACCACTGTGACCTCTGAAAACCTCCTGAACAAAAGGAAGGAGAC
TTGGAAATCCTGAATGGGCTTGGAGACATTAAGGGAGAACTGCCTCCCTGGACCAA
GGCAGAATTCAATAGAACCAGCAAGAAATTTTCCTATGAATGGGAAAGCAGGTGGC
AGGGGGCAGGGGTGGAAAAGCTTTGTACAGGAATTGTGGAAGCTTTTGCATTAT
CTCTAGTCTGAAAGTCACATTTCTCAGTTCCTTTCCACTCTCTTCTGTCAACTTGC
TGTGAGTAAATGACATCTGTACCTGTGACACGGGCCAGGGACTATCACCATATGG
CCCCACACATTATCTAGTACCAGCCTGCCTGGGCCATGCCTTTTCCAGTCACTGT
ACCAGCC

~~Sequence ID - 1214~~ SEQ ID NO: 428

nt:

620

CTCTCCTGTCAACAGCGGCCAGCCTCCCAACTACGAGAATGCTCAAGGAGGAGCAG
GAAGTGGCTATGCTGGGGGCGCCCCACAACCCTGCTCCCCCGACGTCCACCGTGAT
CCACATCCGCAGCGAGACCTCCGTGCCCGACCATGTCGTCTGGTCCCTGTTCAACA
CCCTCTTCATGAACACCTGCTGCCTGGGCTTCATAGCATTCGCCTACTCCGTGAAG
TCTAGGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCAGGCCTATGCCTCCAC
CGCCAAGTGCCTGAACATCTGGGCCCTGATTTTGGGCATCTTCATGACCATTCTGC
TCGTCATCATCCCAGTGTTGGTCGTCCAGGCCAGCGATAGATCAGGAGGCATCAT
TGAGGCCAGGAGCTCTGCCCCGTGACCTGTATCCCACGTACTCTATCTTCCATTCCCT
CGCCCTGCCCCCAGAGGCCAGGAGCTCTGCCCTTGACCTGTATTCCACTTACTCCA
CCTTCCATTCCCTCGCCCTGTCCCCACAGCCGAGTCCTGCATCAGCCCTTTATCCTC
ACACGCTTTTCTACAATGGCATTCAATAAAGTGTATATGTTTCTGGTGCTGCTGTG
ACTT

~~Sequence ID - 1215~~ SEQ ID NO: 429

CACAAGATAGAATGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTTAA
GTGACAGTGCCATAGTTTGGACAGTACCTTTCAATGATTAATTTTAATAGCCTGTG

Marked-Up Copy

AGTCCAAGTAAATGATCACTTTATTTGCTAGGGAGGGAAGTCCTAGGGTGGTTTCA
GTTTCTCCCAGACATACCTAAATTTTACATCAATCCTTTTAAAGAAAATCTGTAT
TTCAAAGAATCTTTCTCTGCAGTAAATCTCGCAGGGGAATTTGCACTATTACACTT
GAAAGTTGTTATTGTTAACCTTTTCGGCAGCTTTTAAATAGGAAAGTTAAACGTTTT
AAACATGGTAGTACTGGAAATTTTACAAGACTTTTACCTAGCACTTAAATATGTAT
AAATGTACATAAAGACAACTAGTAAGCATGACCTGGGGAAATGGTCAGACCTTGT
ATTGTGTTTTTTGGCCTTGAAAGTAGCAAGTGACCAGAATCTGCCATGGCAACAGGC
TTTAAAAAAGACCCTTAAAAAGACACTGTCTCAACTGTGGTGTAGCACCAGCCAG
CTCTCTGTACATTTGCTAGCTTGTAGTTTTCTAAGACTGAGTAACTTCTTATTTT
TAGAAAGTGGAGGTCTGGTTTGTAACCTTTCCTTGTACTTAATTGGGTAAAAGT

~~Sequence ID 1216~~ SEQ ID NO: 430

nt:

484

CAACCTTAGCCAAACCATTTTACCCAAATAAAGTATAGGCGATAGAAATTGAAACCT
GGCGCAATAGATATAGTACCGCAAGGGAAAGATGAAAAATTATAACCAAGCATAAT
ATAGCAAGGACTAACCCTTATACCTTCTGCATAATGAATTAAGTAGAAATAACTTT
GCAAGGAGAGCCAAAGCTAAGACCCCCGAAACCAGACGAGCTACCTAAGAACAGCT
AAAAGAGCACACCCGTCTATGTAGCAAAATAGTGGAAGATTTATAGGTAGAGGCG
ACAAACCTACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAAC
TTTAAATTTGCCACAGAACCCTCTAAATCCCCCTTGTAATTTAACTGTTAGTCCA
AAGAGGAACAGCTCTTTGGACACTAGGAAAAAACCTTGTAAGAGAGAGTAAAAAATT
TAACACCCATAGTAGGCCTAAAAGCAGCCACCAATT

~~Sequence ID 1217~~ SEQ ID NO: 431

GACAGGCGGGGGCCCAGCGGCCGGGTGAAGGCCGGGTGGCTCTGTGAATCAAAGGA
GAGTCCCAGAAAACCTGTGACTGTTGAAGAAAATTCATCTGTGAATTTTTATATTC
AAGGAGTCAGTATTTATATTCATCTTTTAACTGGGAAGATTTATATTTTACTTTA
AACTTCTTGATAATAATTTACAATGAATGGACACAGTGATGAAGAAAGTGTTAGA
AACAGTAGTGGAGAATCAAGGTAAGTAAGCACTTTGTTATCAATTGTTTACTATGA
AGAGAGTTGAAAACCTTGACTTTTTTCTTTATTGTTATTGTTGTTATTTAGTTTTCC
TCATAGGTAGCAGAGTTTTTCAGGTTTTCTCTTAGCTATCCAAATACTAAAAAAT
TCTGATATACGAACCTTTTTTCATAATACAGGTTTTTAATTATATTTTTCATTCAGA
TACACAGTAGATCTTAAATATAGAAAGTTTTTGTTTACTTAAATCTATTTGGAAGT
TTATATTTGAGCTAATAATTAAGCTGGAGCATGTATAATAGATTTAAATTGTTTTG
ACTGTTAGTGAAATTT

~~Sequence ID 1218~~ SEQ ID NO: 432

CTCACTTGGTGGGTGAGCCTCCAATGACTACACCCAAGGAGGATTTAACACAGGGA
TTTTATGACTTGCAACAAGTCAGGAGGACATGGGGTTGGGGTAGTTCAGCAGTGCC
TGTCTGAACAAAGGTGAAAATTGGGCTTTTATTGGGCTGATCAAGGGGGAGTAAAG
GCAGCCAGGAGCAGTCGCCTGTCATGCTTCTACCTATATTGCATGTATAGAAAAGG
GAAAATAAACTCCTTCCTGGGCAGGGTTTTAGTATGCTAAGGAGGGGAGTTATTCA
ACTTCAATCCAATCAAGCATCAGCATTGCTGCGTCCATCCCAGTTTTGTTTTGCT
GGGGCTGAACTTCTTCCTATAACTTTTTGAAACAACAAGAACTCAAGGTGTGACAG
TTACAAGTGGGCCCTTTTTTCACAGTGTGTACCTAAACACGTGAGGACCCTGGATTA
CAGAATGACAGACTCGAAGTGAAGTCAAGTCCGGTTGTTTCATCTTTAGATGGTAAA
GATGGCTGTACGTACTATCCTTGCTTATTTCCAATCTATTGTTTAAACTCTTGTAT
ATGTAATACCGCAGAGGCTAGAGATACAACCTTTGACCAAATGAGTGAATTCAAGT
AATCCATTACTAATGTGATCTGGAAACAAACATGGTGTGAATGTGCATATGT

~~Sequence ID 1219~~ SEQ ID NO: 433

nt:

559

CTTGGCAGCTCCGTTATGTGCCCAGCTCTTTGCAAGGGCATACTGGGAAATGAGTG
GAGATAAAGGACCCAATCATAAGCATTTTACAGTATGGATACCCCATTTTAAAAAG
GTAAACTGAGGCACAATGCAATTTTTTTTTTTTTTTAAGGAGTTTATTTGAGCAAA
CAGTGATTCATGAATCAGGCAGCACCAAACCAGAAGGAGGCTTTGCTGAANAAGGA
TGAGGGACAAGCATTTATAAAGTGAATGTAGATGTAATACAAAGAAAATATTTGAA
CCGGGTGCGGTGGCTTACACTTGTAATCCCAACACTTTGGGAGGCCAAGGCGGGCA
GATCACAAGATCAAGAGATCGAGACCATCCTGGTCAACATGGTGAACCCCATCTN
TACTAAAAAATACAAAAATTANCTGGGCGTGGTGGTGCCTGTAGTCCCAGCT
ACTTGGGCGGCTGAGGCAGGANAAATGCTTGAACCCGGGAGGTGGAGGTTGCAGTA
AGCCGAGATTGCACCATTGCACTACTCCAGCCTGGTGACAGAGAGAGACTCCATC

~~Sequence ID 1220~~ SEQ ID NO: 434

GANNNGTGCGATANNATGNNTGTCTTTTTTTTAAAGTNTTTCNNATNGNAGNGAAN
CCCCCNANNTNNCATAANGAGAGATNACTACNGTACANATAGNGNCANACNGATA
GTAGTANCAANATTGTNTTAGCTANATNANTCAATAGATATCNAGATANAANAANA
NCNNGGATATACAGCGATGTNTNANNGGNNNNNNNANGGAACGAACATCNACNTTA
ANNATAAGCTNGNGGAGAGAGACANGTANGTTATANANNAGAATNGNAGTAGNGT
GATCATAATAGNNNNNANNTANTATATANGATNTTANTGNNCTNTNNTNNGTTTAT
CNNNAATNTCTATNCTNGAGAGNAGCNNNATNNNNAGGCGANGANATTGGGNNNTN
CTCNTNATAGANANCTGGTGTCCNNANAANTACNTCATCTATTNANCTCTCACNANA

Marked-Up Copy

TGGNANNATANAGNAGNGNNNTNNANAGGANTANGCATAGNGNNTNNCTNAAACAA
AANNNATAAGANNTCTCGNNAANANGGGCCTNTNNTNTAGCGAGGNNTTANTTTNT
ATANTTNTTCNCTCTTNNAATANNTANGATANATGANCTNGNNGTGATANATANANN
NNTACNGTNAANNTNTANTCNTATAATAGATANAAATATAGGATNTTNCCTCTGGCN
GGTNGAANANTTNNTNCNNTTTNAATAATGNTGTTAGNGACNGNGNTNTNANANN
NNTTAGAAAGGTACTCTATATACTNNTATGNTNCGGCNNATAATANAACAGATGTT
TGTATNAATATNAAANAAGGTCNNTTTTCGNCAAGAGAANNNTGNCTGGTNATAGAA
TTAGCATAANTTANNTANTATGATNNTANTNNTNCTACNANTNTTAGCNNTTNGCAG
NAGTCATTNNGNATNTATNNGNNTANTAGTNANTTGGGNCTNNTNCAGANTATAT
TNTGNGAANATGAANNTACGNANTCCTNNGNANTATNATNNTGANTANGANAANCN
ANANNTNTTNTANNANTGNCTATANATTGCCNNGATANATTNTNNAATGAANCGA
TAGCCCGCNCTAAGGANNTNNGTNANNTAAANNTCTCAGATAANNTACNTNTTNT
TATTAANCNANNATCACANTATANCNGNGACANNNGCGANANTATATGTATGNNAN
TATNACNGNTCCNNNCCGNGAANNTANTCNTANNAGGCATTCNGNGAGCTNTTCT
NCTAGACNATTTNNANTGAAANNATGCNGNNAACGACNNNCTTNAANTTNTGT
CTACANTCCGCNNTNTTNTTACAGATNGCAGNTAAGNNNANTNANNGCTCTCANCT
NGCTNNNACT

~~Sequence ID - 1221~~SEQ ID NO: 435

nt:

741

AAGCAGAANTNTCTCTAAAAACATTATCTCCTTAAATCTTGAGGTGCATATNAGA
GCCACAGGCAATCTCTGACATATAAAATTGCAGTACAGGCCTTTCAAATTTGGCAT
TTCCTGTTACAATACAACAACCAAGATATATAAATACTGTACAGTGCCTAGACAT
TCCAGTAAGAACCATTATTTTCTTTAATGTAGAATGATTAATACATATTCTACAAG
GGGCAGTAAGGTAGTAATTCTATAGGGTATGTCCCGACATAATTTTCAAATTTGTA
CAATAACACAAACAACTTTGTTAAGGCCATGTTTTATTTGCTGATTAATGGACAAA
AGGCAATGTAATTTATTTTCAAGTATTTTCTTGAAAGTCTGTGCTCATAAAAATCA
TGAAAAGTTGGAAAGACTGTTAAATCACTGAAACTTCAAATATATCTTACACAATC
TTGTTTGTACAAAAATACAAGTTAAATATAAACATAAAGCAATCATGGTAATTTTA
TGCAAATCTGTTTTATGTGATCATCAGTTATATATAAAAGTTTCTCAGTTCTGTTA
TTTGTGAAAAGATCAATACCAGATTGAATGACTACCTATTGGCAAAGGGCCCTAAA
AAGCTTACTTTAGCACTCATCTTTTACATGGTTAAATGCATTTCTTAATTTGAGAT
CACCTAAACACTGGAAAAGAAAAAAATGAAAGGGCAGTATGTCCATAAACCAACA
AATAATTTGGCTG

~~Sequence ID - 1224~~SEQ ID NO: 436

nt:

485

CGAAATTTTCCTTGTGACACAGAGGAAGGGCAAAGGTCTGAGCCCAGAGTTGACGGA
GGGAGTATTTTCAGGGTTCACCTTCAGGGGCTCCCAAAGCGACAAGATCGTTAGGGAG
AGAGGCCAGGGTGGGGACTGGGAATTTAAGGAGAGCTGGGAACGGATCCCTTAGG
TTCAGGAAGCTTCTGTGCAAGCTGCGAGGATGGCTTGGGCCGAAGGGTTGCTCTGC
CCGCCGCGCTAGCTGTGAGCTGAGCAAAGCCCTGGGCTCACAGCACCCCAAAGCC
TGTGGCTTCAGTCCTGCGTCTGCACCACACAATCAAAGGATCGTTTTGTTTTGTT
TTTAAAGAAAGGTGAGATTGGCTTGGTTCTTCATGAGCACATTTGATATAGCTCTT
TTTCTGTTTTTTCCTTGCTCATTTTCGTTTTGGGGAAGAAATCTGTACTGTATTGGGA
TTGTAAAGAACATCTCTGCACTCAGACAGTTTACAGA

~~Sequence ID 1226~~SEQ ID NO: 437

GGTTTTTATACTTGCCATGAAACTGTTCTTTGGGATATTATTTTGTTCAGGTTCCC
CACTTGGACAGCAGAGGGGGTGACTCTGCCCATCCCTGCCACTGGTAGCCAGGCGG
GCAATGTCTGCTAGCAGTCTGCTTCTGTCTGAACTCAGCCAGCAGAGGCAAACCTCC
CGGTTCCCCGAGAAACACTCTGAAGGCAGGGTGGGTGACTCCACCCACCACCGCCT
CTCCTAGCCATGCAGGCCATGTCTGCTAGAGCTTCCAGCGCAGTGGTCCTAATTCT
GTCTGAATCCGGCTGAGGGGTGCAGCCTCCTGTTACTGCCCAGGGAAACACCCAGA
TGGCAGGGTGGGTGACTCCAACCACCTCTGCCTGTGGTAGCCAGATGGGCCACACC
TGCTAGAGCTTCCAGCCCAGCAGTCCCGCTACTCTGTGGGTGGGTGCCATCCCCTG
TTCCTCTGGGAAGCACCCAGACAGCTGATTACGTGACCCCAACCACTTCTGCAGAT
CCTAGCTGAGCAGGACTTGCTGGTTTGGACAATGCCCAAGCAGGGAAGAGCCCTCA
TTCTCTTATCACTGACAGAGGTGAGATGTCCGANTTTGTANGCTGGTGGAGGAGTG
AGGTGGAGGAGGTATGCCTCT

~~Sequence ID 1228~~SEQ ID NO: 438

GTTATTCAGGTATCCATCAAAATTTTATAAGAGGGCCGGAACATCGGCTCACACC
TGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGTTCACCTGAGGTCAGGAGTT
CGAGACCAGCCTGGCCAACATGGCAAAACCCCGTCACTATTAAAAATACAAACAT
TAGCTGGGTGTAGTGGCAGGTGCCTGTAATCCCAGCTATTTCGGGAGGCCTAGGAAG
GAAAATGGCTTGAACCTGGGGGTGGAGGTGGAGTGAGGCAAGATCACACCACTGC
ACTCCAGCCTGGGCGACAGAGCGAGACTCCATCTCAAAGAAGAAAAAAAAAACAA
CAAAAAAACCTTTATCAGATTATCAGAGGTTATCACTACAGAGGGAGGTAAAATTG
GAGGGAAAAGGGTACAAATTTATTTTAC

~~Sequence ID 1230~~SEQ ID NO: 439

nt:

741

AAGCAGAANTNTCTCTAAAAACATTATCTCCTTAAAATCTTGAGGTGCATATNAGA
GCCACAGGCAATCTCTGACATATAAAATTGCAGTACAGGCCTTTCAAATTTGGCAT
TTCCTGGTACAATACAACAACCAAGATATATAATAACTGTACAGTGCCTAGACAT
TCCAGTAAGAACCATTATTTTCTTTAATGTAGAATGATTAATACATATTCTACAAG
GGGACAGTAAGGTAGTAATTCTATAGGGTATGTCCCGACATAATTTTCAAATTGTA
CAATAACACAAACAACTTTGTTAAGGCCATGTTTTATTTGCTGATTAATGGACAAA
AGGCAATGTAATTTATTTTCAAGTATTTTCTTGAAAGTCTGTGCTCATAAAAATCA
TGAAAAGTTGGAAAGACTGTTAAATCACTGAAACTTCAAATATATCTTACACAATC
TTGTTTGTACAAAAATACAAGTTAAATATAAACATAAAGCAATCATGGTAATTTTA
TGCAAATCTGTTTTATGTGATCATCAGTTATATATAAAAGTTTCTCAGTTCTGTTA
TTTGTGAAAAGATCAATACCAGATTGAATGACTACCTATTGGCAAAGGGCCCTAAA
AAGCTTACTTTAGCACTCATCTTTTACATGGTTAAATGCATTTCTTAATTTGAGAT
CACCTAAACACTGGAAAAGAAAAAAATGAAAGGGCAGTATGTCCATAAACCAACA
AATAATTTGGCTG

~~Sequence ID 1231~~ SEQ ID NO: 440

nt:

203

TTGAGGAAGGGTCTACTGTCTTTTTTAAATGGCACAATTTTAAGAGGTTTGAGAGGT
ACAGTCCCTTAACCTGCCACGGGAGAGGGGCCCCCAAACCTTTCTTCCCCCACA
TCTGGTTTTCTGTGTGGAGGGGGAGCAGGGATATCTAAGCTGTGGTGTGAAAGGGT
AGGAGAGATGCTGGAGGTGGGGGTGCTGTGTTCTA

~~Sequence ID 1239~~ SEQ ID NO: 441

TTTCCTCGGGAAGCGCGCCATTGTGTTGGTACCCGGGAATTCGCGGCCGCGTCGAC
ATTTTTTTTTTTTTTTTTTTTTTTAGAAATGATTAACAATTTATTGAGTTTTATTTATC
TACAAAAATATAGCAATACAGNGAACTTCACCAAATCCTAAATATTAGTACCTGA
ACTGGCTACAACACCGNGTGCACACCCAGTTCCTGCAGAATCTCTTGCAGATATGG
GAGAGTCAGCCAGTGAAAAGATCCATTTCTTGGGAATCCTTGTCAACAAGACCAGT
TCAGAAATCCAGGATATATAGAAGCCTACTGTAATTTAAAAACAGTAACAAAAACC
CCAACAAAACCCAAATCAACAAAGACCAAGATAAAGNGTGATAAACATTAATTGT
AATGGTTTTCTTTACATGCAATACATGCATTTTAAATCACTAAGAAACACGAAA
TTTTGTAGAGCAAAGTTTNGTTTTCACGTAAGTGCAAATGAATATATATTTTATTT
TTTATACTATTAAATTATATATATTTTTTTCCATACAAAAGCACACAGTGTTAATCT
ATAAAATGACATCCAAGTGGATGATGATTGTTTTTGCATGTCCCCCTGCTTAGATT
TTTTTAAATATATAGTCAAAAATTAACATCCTTCTTTAAAAATACAGAAGGGAAA

Marked-Up Copy

AANGGGCAAAAAAAAAAATCTAGACTCGAGCAAGCTTATGCATGCATGCGGCCGCA
ATTCGANCTCGGNCGACTTGGCCAATTCGCCCTATAGNGAGTCGNATTACAATTCA
CTGGGCCGNCGNNTTACAACGTCGNGACTGGGAAAACCCTGGCGTTACCCNNCTNA
TCGNCTTGNAACAATNCCNNTTNGCCAGNNGGG

~~Sequence ID 1255~~SEQ ID NO: 442

TCACTTCGTATNGAANCTGTTTGGACTTGCTCAACAAGACCTTATCTTAACAAAA
GTAAC TTATAGAAAAGGGAGACATTCATTTAACTTCAAGCCCATATTATTCTTAAA
AGCTGACTCTTGAAATAGTATTTATTGAGTCATAGTGGAGTCATGGGACTTTTTAA
GGGCCGGAAGGGACTATTTAGATCATCCAGTCCCACCCTGTCATTTTATGGAGGAG
GAAACTGAGGCCTAGATAAGATAACCAGTTAGTGGGTCCACTGACCTTTAGGACAG
TAGTCTATCCGTAAGAGACAACATGGAGAAAGAAATACAACGTTTTTATAGTGAAT
TATCATCTTACAAAGAATATTCTTCCCATATCGCACTTTTAAAAAGTGGGTACCTT
AGTCAAATAGGAGAAAAAACCACTTGAGTAGTTTCATCCTCAGGTTTTAGGTGAGG
AAACTGATACTCAGATTAAATAACTTTAAGCACACAGAGCCTGAATGATAGTCTTA
TTTGAGCTCATCTGTGCTTTTAATGTGTACTACGTTAGGTGTTTTCACTTGCAATTT
CCTTTAGTCTTATTTGAGCTCATCTGTGCTTTTAATGTGTACTACGTTAGGTGTTT
TCACTTGCAATTTCTTGTGTTGACGTTGACAATAAATCGTGAAGCTGCCTTATCTAA
GNAGTCCTAAAGTAAATCATTGGAACACATGTANCCAGTTTGTTGTTTTTTATTTGC
CAGGTNTCAAATATAACTGAAAACCCATGCTAACTGACTNATTTTAAAAGNTGTNT
GGGGCATGAAANGATTGCTCTGCCTGGGCGGGNGGTTNANCCTGNGTCCCCCNNTT
NGGAGNCCACCCANGANGCGATATTTNAGGGNNGATTCNAAACCCCTGGCACGNNG
NAACCCCNNTTTTAAANANAAAAANANCGGNNG

~~Sequence 1256~~SEQ ID NO: 443

TTGTGTTGGTACCCGGGAATTCGCGGCCGCGTCGACGGAGTTTTACCTTATTACAC
TTTAATCTCTGGATTTACCCCATCTCATTCTCTTTTAGGAAAAGTGGTTGTATGT
GGTGGCTTTGATGGTTCTCATGCCATCAGTTGTGTGGAAATGTATGATCCAAC TAG
AAATGAATGGAAGATGATGGGAAATATGACTTCACCAAGGAGCAATGCTGGGATTG
CAACTGTAGGGAACACCATTTATGCAGTGGGAGGATTCGATGGCAATGAATTTCTG
AATACGGTGGAAGTCTATAACCTTGAGTCAAATGAATGGAGCCCCTATACAAAGAT
TTTCCAGTTTTTAACAAATTTAAGACCCTCTCAAAC TAACAGGCTTAGTGATGTAAT
TATGGTTAGCAGAGGTACACTTGTGAATAAAGAGGGTGGGTGGGTATAGATGTTGC
TAACAGCAACACAAAGCTTTTGCATATTGCATACTATTAAACATGCTGTACATACT
TTTTGGGTTTATTTGGAAAGGAATGCAAAGATGAAGGTCTGTTTTGTGTACTTTTA
AGACTTTGGTTATTTTACTTTTTTGAAAAGAATAAACCAAGAATTGATTGGGCACA

Marked-Up Copy

TCATTTCAAGAAGTCCCCCTCTCCTCCACATTTGTTTTGCCAATTTGCACATTAAAT
GACTCTTCCCTCAAATGTGTACTATGGGGTAAAAGGGGTAGGGNTTAAANATGTAA
ACAGTTGGGTTTTTTAAGGGNCCTTTTTTCATAACTGGAACACTCTNTACAAGGNTN
CTTNTTAAATAAAATAACTTGACTTTTTTGTNTTAAANGNANCTTCNTGCTTCCA
TAAAAAAAAAAATTTAANTNGNCANCTNTGCTGCTGCGNCCANTTNGCTNGNCCNT
GGCATTCCCTAGGGANGNTNAATANTGGCNNNTTAACNNGGCNGNAACNNNNNCCA
NT

~~Sequence ID 1331~~SEQ ID NO: 444

GGGCGATGCATGCTTTATTAAGGCTCTTGTTTCACCTGGCAGTGTACTGTATCAAC
GTATAATACAGAAAAAAAAATCTCTTTAAGGTCCTCCTTCACAAAGACATAGAGTGA
AACTCCCTTTACATGTCAGTATTTGTTCAACACTTTAGGCAACTTGACTGTCAGTG
TTAAAATGGAAAACAGGAAAATGGAAAATCTGACCAATTCTGCCACCTTGAGACT
TTCATATAGACCTTGCACAACAATTGTATAGATCACACACCGGCTGTATTTAATAT
GTAACATTTTCACACATATTAAAGATACAGAAGTATTAAAAAACCCCAATGTTAA
TGTATTTGCTTAAAAGGCACAAGTTTCACATATCTGTCTAGCTATCTGTTGGTAAT
ACAGAAAGTATACTACTTTTTTAAAAAGTGGGCAGAATTCTTGTGTATGTATATT
TGTGTGTACAGTATGTGTATGTGTGTATATATATATATATATATAGATAATAT
ATAAATATTTTTTTTAAAGGAGAACTAGAATGTTTAGCTAGAAAATTCACAGCCT
GTGAAGAAATATTTCAAATGGCCATAAAGGAGGTAAAAATGAAAACCATAACCTA
ACTTTTATAGAGGCTTTATCTTTAATTTAACGATGTGCGGAGGACTTTCTTGCTTG
AATCTGTTCCGGGCTGTCTGCTCTGTCCATCAAATGGGCAGGTCTGGGAATGAGGC
ACCTTCGGCCGTTTCAAGAAGTGGCCTGAACAGAATGCTGGAACCCAGGCTGGACTCG
GAC

~~Sequence ID 1332~~SEQ ID NO: 445

CAAACCTGCATGTTCTGCACATGTATCCAGGAACCTAAAAAAAAAAAAAGATAGTT
TGTGTGTCTTAATTGAATAATAGTAGATTTATAGATTAAAGATCTATGGGTTTTTA
ATATGGATTAGAAATCTGTGGGTTTTTGATATGGATTAGAAATCTGTGGGTTTTTA
ATATGGATTGGAAATCTGTGGGTTTTTAATATGGATTAAAAACATCTGTGGGTTT
TTAATATGGATTAAACATCTGTGGGTTTTTAATATGGATTAAACATCTGGGTTTTT
AATATGGATTAAACATCTGTGGGTTTTTAATATGGGTAAAAATCAAAGAAAATG
AACTATTTGCTCCAGTGCAGGAAAATACAGGCAATACTGGATACAATTAGATGGTC
AGGAGCGATAACCCGGTTGCCATTGTTTGAAGAAGAGAATAAGGTGCTAGCATTCC
TATCCGTAGATAATTTGACAGCTAGGAAATAGGGGGAGTCTTCTATGTAGTTAGTG
AAGGCTAAATGAACTATTATATGCAGTTATCGTAGAAGAGTACTCAAAAAATCTG

Marked-Up Copy

TAAAAAATAAAGAAAGGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTG
GGAGGCCGAGGCGGGTGGATCATGAGGTCAGGAGATCGAGACCATCCTGGCTACCA
NGGTGAAACCCCCGTCT

~~Sequence ID 1335~~SEQ ID NO: 446

CAAGACTCCATCTCAAAAAAAAAAAAAAAAAATCTACAGTGCTGAGTATATAAAATTAT
TAACACATTTTACAACAATATGTGTTTGTGGAGTTAAATATTTTTTGTCTTTAAAA
CAGGTAATTTTAGTGCACTTAATTTGATGATTAAATATGGTAGAATTAAGCATT
TTAAATGTTAATGTTTGTACATTGTTCAAGAAATAAGTAGAAATATATTCCTTTG
TTTTTTATTTAAATTTTTGTTCCTCTGTAAACTAAAAGAACACGAAGTAATTGGTC
ACAATTACTGGTGTTTAACTGCCAAATATGGGTAAATAAGGGAAAAATTTGTTTTAA
TATTTAGTCCTTCTGAGATGGCTTGAATATTTGAATTTTGTGTACGTCTATACTG
GGTAGTCACAAGTCTTATAAACACTTTAGAGGAAAGATGGATTTTCAGTCTGTATTT
TTAAACATCATTTATTTTAAATCTGGTGCTGAAAAATAAGAAAAAAATTTAACTGC
ATTCTGCTGTTCTTCTTTAGAAAGCATTCCTGCGTAAATACTGCTGTAATACTGTCA
TGCAAAGTGTATCCTTTCTTGTCTATCCTTTTTTGGGGCAGTGTTTTTTTTGTTTTT
TTCCTAGAAATGTTTGTCTTCCCCACCTGTTGATCCAGGTTAAGGAATACTTTT
TTACACTTTATTCAA

~~Sequence ID 1336~~SEQ ID NO: 447

CTTTTCCTCCCGCTGTCCCCACGGAGGGGACTGCTCTCCCCGCTGCATCCTTTC
TGTGAGGTACCTTACCCACCTCAGCACCTGAGAGGGTGAAATAGAATTCTAACCTC
GACATTCGGAAGTGTTTTTGAGAAGTCTCGGTTCGGTAAGGGAAGTCTTCCAAGTC
CGTGCAGCACTAACGTATTGGCACCTGCCTCCTCTTCGGCCACCCCCCAGATGAGG
CAGCTGTGACTGTGTCAAGGGAAGCCACGACTCTGACCATAGTCTTCTCTCAGCTT
CCACTGCCGTCTCCACAGGAAACCCAGAAGTTCTGTGAACAAGTCCATGCTGCCAT
CAAGGCATTTATTGCAGTGTACTATTTGCTTCCAAAGGATCAGGCCCTGAGAACAA
TGACCTTATTTTCTACAACAGTGTCTGGGTGCGTGCCAGCAGATGCCTCAGATAC
CAAGAGATAACAAAGCTGCAGCTCTTTTGATGCTGACCAAGAATGTGGATTTTGTG
AAGGATGCACATGAAGAAATGGAGCAGGCTGTGGAAGAATGTGACCCTTACTCTGG
CCTCTTGAATGATACTGAGGAGAACAACTCTGACAACCACAATCATGAGGATGATG
TGTTGGGGTTTCCCAGCAATCAGGACTTGTATTGGTCAGAGGACGATCAAGAGCTC
ATAATCCCATGCCTTGCGCTGGTGAGAGCATCCAAAGCCTGCCTGAAGAAAA

~~Sequence ID 1337~~SEQ ID NO: 448

CAAGAACTCTGGGACATTTGCAAAGGGTATGGCATATGTGTAATGGGAATACCAGA

Marked-Up Copy

GGAGAGGAAAGACAGGAAGTCAAAAAAGAATTTTTCCAAATTAATGATAGGTTCC
AAACCACAGATGCAGGAAGCTTAAACACCAACAGGATAAATAAAACAAAATCTACG
CTTAAGCATATCATACTTAACCTGCAGAAAATTACAGACAAAGAAAAACACCAGA
GGGAAGCTGGCAGAAACATACCACCTATAGCGGAAGAAGAATAAGAATTACATCA
GACTTCCCTTCAGAAATCTTGCAAACAAAAAGATGTAGCACAATATTTAAAGTATT
AAAGGAGGCCGGGCCCGGTGGCTCGGGCCTGTAATCCTAACACTTTGGGAGGCTGA
GGCAGGAGGACCATGAGGTCAGGAGATCGAGACCATCCTGGTGATGGTGATACCCC
ATCTCTACTAAAAATACAAAAATTAACCGGGCATGGTGACACGCACCTGTAATCC
CAGTACTTTGGGAGGCTGAAGCAGGAGAATCGTTTGAGCCCAGGAGGTGGAGGTTG
CAGTGAGCCGAGATCACATCACTGCACGCCTGGGCAACAGAGCGAGACTCCATCTC
AAAAAA

~~Sequence ID 1338~~SEQ ID NO: 449

CGACCCGTTTTAGTCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCTGCCTC
GGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCGTAAATCAG
GTTTTTTAAATGTTTGCCAAACCTTATCACTGACTTTTATAACAAAATTATTTACT
ATAATCATTAGGGAATATTTAAGTTCTGCTAATACTTAAAATTGCAGAGTGCTAAA
ACCAGCAGTGAGTTTAGAATCAAGCTAAGCTTTTATTGTTGCTACTATTTGAGGCAT
ATTAGTTGACTGGTGTTTCATATGCAAGGCAGTCTACTGGGTGCAACAAGGGTTAGA
AGGATATTTTTTAAAAAACTGACCCTATTCTCAGGATGAAAATAATACACTAGTAAT
AGTCTGCTCTGTTGGTAACTCCTCGTAAGGAGGTACAATTAAAATGCTGTAGTGT
TGCAAGGGAAGGAGAGGAAGAATCATATTCCTTCACTAGCAGGATCAAGAAAGCTT
TTATAGAAATATACAAAATCTTCACTTCTTGAAGGATTGGTAAAATTTAATAGCCA
ACATTGGGCACTTATTCATTCTCTGAGTAAATATTTATTGCATGCTTATCTTGTAT
CAAGCATTGTGATGAAAGCACAGAATGAAAGAGGAGGGAGAATGTTTAGAGAATA
AGGGCTGAAACACAGATTTTGTAGGGAGCGTAGGGGAGACTGANAAGACAGGTTCA
GGTTAGTAAGGGCGCTCATATTTTGACCCTGAATGTTAAGTATGTGCACATCATGC
TAGCTATTCTAAATCAGGCATTTTCAAATGGAAGCAGGCACTGACATTTT

~~Sequence ID 1344~~SEQ ID NO: 450

CGTGAAGGGTCTTTATGTATTAGTATTAGAGTGATCTTTTGATTATTTTCCTCACT
ATAAGGAAATTATTTCTCAGGATGAGCTGCCATAACATTCCACTGTCTGATGGCA
ATTTTAAAGCCTGAAATTGAAGCCCATGGCTAGGCTATGAGAACCCTAGTTCGTAT
AGTAAAGTTGATATCTTCTGGATGTATACTAATTTTAGGCTTTATTTTAAACTGC
TGGAAGCTGAACTTAGACAAAAGTATTTTCAGGACATCATTTACAATGTTTAGCC
CTAAAGAGTCAAGCTGTGGGATTCTGAGTCTTTCATATGTTACAGCAGAACTTAA

Marked-Up Copy

AAGCAAGAGGAAATTGGCTGGGCACAGTGGCTCTGTAATCCCAGCACTTTGGGAGG
CTGAGGTGGGTGGATCATGAGGTCAAGAGATTGAGACCATCCTAGCCAACATGGTG
AAACCCCATCTCTACTAAAAATACAAAAATTAGCTGGGCGTGGTGGCACACGCCTG
TAATCCCAGCTAGTCAGGAGGCTGAGGCAGGAGAATATCTTGAACCTGGGAGGCAG
AGGTTGCAGTGAGCCAAGATTACATCACTGCACTCCAGCCTGGTGACAGAGCGAGA
CTCCGACT

~~Sequence ID 1348~~ SEQ ID NO: 451

CTGAAACTGCACTGAACCCACAGGTAGGTTACATCACAGGACAGAAATCTGAGGAG
CTGGAGAAAGCAAAAGAATAAAGGATGGGCTGACACCAGAAGGAATTAAAGGAATT
TTTATACTGAACTTCAATTACTTGTTCATTTGAAGTTTGTTTTTTTAATGAACGTT
TTTGCTGTTACTTAAATATAGTGTTTTGAAAGTGTTTCAAATGTATTCAAGTTGGG
ATTTTCCATATTTTACTACAGTTCTGTCTTAGTATGTTCAACATAAAACACTTATC
ATTAAAGCTCACAAAGTGCTTTTTTGTAATATGAGGATAAAATGAAGCCATATAAG
AATTTTTTTATATCTGTACATTTAACCCACATTTGAGCTTTAGCCAAAATATATAG
CTTTTTTTTTTCTGACCTGGCCAACGTATTATCCAGCAAACATCAACTGAAGCAAT
ATGGAAACACTTCCAAATGTTTGCCAATAATGCTATTAAGTGACTGATGTCAACAT
TAGTTACATGGCAAACATAAGAGGCATTATACATTTTTTAAAACACACTAACATATA
ACTGTAGATAATGTAAGGTTTATTTATATGCATATTTCATAGTATATTTAAATGTT
TAAATATAAAAAAGGGTTTTTAAACACTTTTAATTTTTTATCTTTGATTTTTTTTAT
TGATATCTCTTTCCAGGCTACTAATAAAATTGCCAGAACTAACTATCAGGTAAAG
GTTAAGGCATCAATTGACAAGTAAGTTTTCTAATTTTCGTTTTGAATTACAATTCCA
AATGTAAGACTTTTTAAAATGAATGGCCTTTATTTTATAGAATAATTTTGACCTTT
TAAATTTACTTATCTAACATTATATAATGAATGTACTTCAAATATTTGACTTTGAA
GTCAACATTAACAAATTCATGGATCCTAATTAATAATTTACTATAAAACTGGAATCA
TTTATTACTTCCTT

~~Sequence ID 1351~~ SEQ ID NO: 452

TTTTTTTTTTTTTAAAAGAGATGGGTCTCACTATGTTGCCCATAAATGTTTATGAG
ATTAAGTTCATCTTTTTTATCTGAGTAGTATTTTATTGTATGAATATACCACCAT
TATTTATCTGTTGGTTATTTCCAGTTTTTGGGCTATAATCCAAAATGCTTTTTTCAA
ACAATAGGCTATATATCATTAATGTCCGTTTATCAGCAGTATAAAATATCTTACCA
TAAATATTAATAAAAGAAGCATTCATATATAAAATATAGATATTTCAAACCCTACA
GAGGGCCTTTTAATGATTAAATATTTTGTCTTACAAAAAGGTCCAGGTAATTACA
CCCATGAGGTTAACCTGCCTTAGTGCAGGACTTAAATAAGGCTTCTCCTGCCATC
TCTCTCCATTTGTAGAATGTGAAATCTTTAAAATGCATCCTATATTAGGAATACT

Marked-Up Copy

ATAGCTGTGCACTGGTGTGTTTGTCTCTTCTTTAAACTCGGGACCGTATATATCTGC
TCAAATTGCCCAAGTATACATATGCTGCACTCCATCAAGTGTCAGGCCACATTCTA
TCAGCACAGCGTGA CTGCCTATCAGTGACAATATAAGTGAGCTCTATTTGGATCCC
TCTTACCCTACCTTTTATATTTATGACAGCATTATCATAAAACTCCAATATTCTTC
AATAACTTACATGTTTGTGTTAGGATAAAATTATTACCCTCAATGAACTACAT

~~Sequence ID 1352~~SEQ ID NO: 453

ACCAGCTTCTTCACAGGTTCCACGAGTCATGTCAACACAGCGTGTGCTAACACAT
CAACACAGACAATGGGTCCACGTCTGTCAGCTGCAGCCGCTGCAGCTACTCCTGCT
GTCCGCACCGTTCCACAGTATAAATATGCTGCAGGAGTTCGCAATCCTCAGCAACA
TCTTAATGCACAGCCACAAGTTACAATGCAACAGCCTGCTGTTTCATGTACAAGGTC
AGGAACCTTTGACTGCTTCCATGTTGGCATCTGCCCCCTCCTCAAGAGCAAAAGCAA
ATGTTGGGTGAACGGCTGTTTCTTATTCAAGCCATGCACCCTACTCTTGCTGG
TAAAATCACTGGCATGTTGTTGGAGATTGATAATTCAGAACTTCTTCATATGCTCG
AGTCTCCAGAGTCACTCCGTTCTAAGGTTGATGAAGCTGTAGCTGTACTACAAGCC
CACCAAGCTAAAGAGGCTGCCCAGAAAGCAGTTAACAGTGCCACCGGTGTTCCAAC
TGTTTAAAATTGATCAGGGACCATGAAAAGAACTTGTGCTTCACCGAAGAAAAAT
ATCTAAACATCGAAAACTTAAATATTATGGAAAAAAACATTGCAAAATATAAAA
TAAATAAAAAAAGGAAAGGAACTTTGAACCTTATGTACCGAGCAAATGCCAGGTC
TAGCAAACATAATGCTAGTCCTAGATTACTTATTGATTTAAAA

~~Sequence ID 1353~~SEQ ID NO: 454

ACATTCTGAAAAGGCAAAAGGGAGGAAGA ACTGATTAGTGGTTAGCCCAGGGTTA
GAGTTGGGGAGAGGATATAATGAGGGA ACTTTTGTGGATTCTGTACCATGATTATG
ATTACACAAACCTATGCATACATTGAAACACATAGAACTATACGTTGAAAAAAGTG
AATCTGCCTGTATGTAAATTTAAAAGAAAAATATTTTTTTAAAAAAACAGATGCTT
CTTAACACATTATCATCTATGTCAGTTTAAACAGTTAGTAGACTTAGGCCAGGTGTC
ATGGCTCACTCCTGTAATCCCAGTGCTTTGGGAGTCTGAGGTGGGACGATCTCTTG
AGACTAGGAGGGAGTTTGAGACAAACCTAGGCAATGTAATGAGACTCTTTCTCTAC
AAAAAATTTTAAAGTTATCTGGACATGGTGGTGCCTGCCTGTAGTCCCAGCTACTT
GGGAGGCTGAGGTGGGAGGATTCTTGAGCCCAGAAGTTCAAGGCTACAGTGTGCT
ATGATAGAGCCACTGCACTCCAGCCTGGGCAACCAGGTGAGACCTTGTCTCTAAAA
TGAATAAATAAAT

~~Sequence ID 1355~~SEQ ID NO: 455

TGGTCTTTCACCCAGCCAGGGAGAAGGTTCTTCGCTCAGTATGAAGAAAAGCAACC

Marked-Up Copy

CAAAACTCTCAATCTGATTTGTTTTGTTTATGTCGATGCCCTGTAGTTTGAAAGT
GAAGTAAAGATTTAGAAATTCACCTAAGTCCAAAGGAAAACACGTGGTTTTTAAAGC
CATTAGGTAAAAAAGTTCTCAATAAAGGCATTACAATTTTTTAGGTTTAGAAAGA
TGGACTTTTCTGATAAATCTTGGCAGACATCTAAAAAAAACCATATTTTTTCACA
AGAAAATGCAAGTTACTTTTTTTGGAAATAATACTCACTGATTATGGATAAAATGG
AATATTTTCAGATACTATATTGGCTGTTTCAAATAGTACTATTCTTTAACTTGT
AATTTTTGCTAAGTTATTTGTCTTTGTTGTATCTATAAATATGTAAAAAATATTTA
AATAGATGTACCTGTTTTGCTTTTCACACTTAATAAAAAATTTTTTTTTTGT

~~Sequence ID 1359~~ SEQ ID NO: 456

CGGGATCCCTAGTATAACACATTCAGTGTTCCTTTTTCAGTCTTACTACTTTGACC
GCGATGATGTGGCTTTGAAGAACTTTGCCAAATACTTTCTTCACCAATCTCATGAG
GAGAGGGAACATGCTGAGAACTGATGAAGCTGCAGAACCAACGAGGTGGCCGAAT
CTTCCTTCAGGATATCAAGAAACCAGACTGTGATGACTGGGAGAGCGGGCTGAATG
CAATGGAGTGTGCATTACATTTGGAAAAATGTGAATCAGTCACTACTGGAAGTGC
ACAAACTGGCCACTGACAAAAATGACCCCATGTGAGTATTGGAACCCAGGAAAT
AAATGGAGGAAATCATTTGCCTTAGGGATTGGGAAAGCTGCCCACTAACTGTCTTC
CCCATTTGTTTTGCAGTTGTGTGACTTCATTGAGACACATTACCTGAATGAGCAGGT
GAAAGCCATCAAAGAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGGAGCGC
CCGAATCTGGCTTGGCGGAATATCTCTTTGACAAGCACACCCTGGGAGACAGTGAT
AATGAAAGCTAAGCCTCGGGCTAATTTCCCATAGCCGTGGGGTGACTTCCCTGGT
CACCAAGGCAGTGCATGCATGTTGGGGTTTCCTTTACCTTTTCTATAAGTTGTACC
AAAACATCCACTTAAGTTCTTTGATTTGTACCATTCTTCAAATAAAGAAATTTGG
TACC

~~Sequence ID 1360~~ SEQ ID NO: 457

TGCGCAGACCAGACTTCGCTCGTACTCGTGCGCCTCGCTTCGCTTTTCTCCGCAA
CCATGTCTGACAAACCCGATATGGCTGAGATCGAGAAATTCGATAAGTCGAAACTG
AAGAAGACAGAGACGCAAGAGAAAAATCCACTGCCTTCAAAGAAACGATTGAACA
GGAGAAGCAAGCAGGCGAATCGTAATGAGGCGTGCGCCGCCAATATGCACTGTACA
TTCCACAAGCATTGCTTCTTATTTTACTTCTTTTAGCTGTTTAACTTTGTAAGAT
GCAAAGAGGTTGGATCAAGTTTAAATGACTGTGCTGCCCTTTACATCAAAGAAC
TACTGACAACGAAGGCCGCGCCTGCCTTTCCCATCTGTCTATCTATCTGGCTGGCA
GGGAAGGAAAGAACTTGCATGTTGGTGAAGGAAGAAGTGGGGTGGAAGAAGTGGGG
TGGGACGACAGTGAAAT

~~Sequence ID 1361~~SEQ ID NO: 458

TATAAATACACTCCGGGATGATTTACCCCCGGAGGTCAGCTAGTAAAATACATGAG
TAGAATTCCTTAAAGTATGTGATAATTGCTCATCACTATCCAAGTGTGACATAAAT
CATAAAAAGAATTGACAAAATCAGGGTCGCAAAGAGAATTGAAAAAATCTGTCAC
AACCAAAATTTAAATTGACCTCTGTCCTAGAGTATGAGAGCCACACTGAACAGAAA
AACCAGATAAATCTTTTATAAAATATTCATTTGCAGCCCCATTAACGTTGCTTGTC
ACCCACCTCCCCATGTCCTTGGACAACTGAATGTATAGTAACATCATCCCAGGC
CAGGCGCGGTGGCTCATGCCGTGAATCCCAGCACTTTGTGAGGCTAAGGCAGGCAG
ATCAGGAGGTCAGGAGTTCAGGACCAGCCTGGCCAAAAAGGTGAAACTCCGTCTCT
ACTAACAATACAAAAATTAGCTGGGTGCGGTAGTAGGCGCCTGTAATCCCAGCTAC
TCGGGAGGCTGAGGCAGGAGAATTGCTCAAACCCGGAAGGTGGAGGTTGCAGTGAG
CTGAGATCGTGCCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCTGTCTCGGG
GAGGGGGGTGGCGGAGATAAAGAAATAACATCATCTTATACTGTCAAGCTCAAGGT
GTCTGCAGCCTTATCTTCAGGGGAAGTTGTGTCTTTCTCAGGGAAGATACAGATTT
CAATTTAGAGCAAGACAGAGAGAAGTTACATTCAGAGAGGAAAATGCAGTAGTCTA
ACTG

~~Sequence ID 1364~~SEQ ID NO: 459

GCGGCCGCGCTCTTTTCAATTTTTTAAAAAGAAGTTTGTTTTCCATTTTCAGTAATTT
CTGCTTTGATCTTCCTTATGTCCTCCTATTGAGTTGATCAGCTTTCTTTATTCTTG
CCTTTTCTCCTCTGTGTGCCCTTTCTATTAACGTATTTACCCTTAGGCTGGGCACA
ATGGCTGATGCCGTGAATCCCTGCACTTTGGGAGGCCGAGGCAGGTGGATCACCTA
AGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTGGTCTCTACTAAA
AACACAAAAATTAGCCAGGCATGGTGGTGTGCACCTGTAATCCCAGCTACTCAGGA
GGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGCGGAGATTGTGCCAAAGCACTC
CAGCCTGGGCAACAAAATGAGACTTTGTGTCT

~~Sequence ID 1365~~SEQ ID NO: 460

CACCAGGCTGTCTTCAGATACTTCATACAGAAATGAGCCTCCCTGTGGGGTCCTCT
TCCCTCCTTCAGCCTGTCCATCAACACAGCATTGCGGGATCCTTACCATGGCATCC
AGCCCTGGAGATGCTTCAGGAAAGTTGCAGGTCCATGCTGCAGGACAGGCTCAGAT
CAGCAGAGACGCATCTCACATCGGGCTGTGAAATTCAAGTTGAGCTGCAATTGGCA
ATGAGAA

~~Sequence ID 1366~~SEQ ID NO: 461

Marked-Up Copy

GTTATTCACTGAGACCGTGCCCCGGTTATGAGGTTGTACCAGAAAGCAAGTATTCA
CTATGCACACTATTACCGCTCACCCCTAGCATTGAAGCCAGCCTGTAGCCTGAAAG
CCTTTGCTTTGAGGGCAGGTCTTTCCCCAAAATGCAGACACGAAGGTGCAAAGTGA
AGCTGCCAGTCTTGCAAAGATGTAAGTTGTACGAAGGCCACGAGTGGCAGGGAG
AGCTGTCCCACATTTGCGGAAGTGGCTATGTGAGGACGGGGGAGGCGGGTCCCTTA
GAGATGAGACAATCATAAGGGGAGATATCAGAGAAAATCGTAAGGGGAGCAGATGG
TTGTCAAGAGAATAGGCTGACCATCGAAGGACTGGCAGAAGCTTTCAGAAAACCAC
TGGACGGCTGGGCACAGTGGCTTAGGCCTGTAATCCCAGCACTTTGGGAGGCTGAC
GCAGGTGAATCACTTGAGGTCAGGAGTTCCAGACCAGCCTGGCCAACATGGTGAAA
CCCCATCTCTACAGAAAATATAAAAATTAGCCAGGCGTGGTGGCACAAGCCTAGAA
TCCCAGCTACTTGGGAGGCTGAGGCAGGCGAATGGCTTGAACCCAGGAGTCAGAGG
CTGCAGTGAGTCGAGATTGTTCCACTGCACTCCAGCCTGGGTGACAGTGCAAGACT
CCTTCCAAAAA

~~Sequence ID 1367~~SEQ ID NO: 462

TTCGTGAGTGATGGCGTCCCGGGTTGCTTGCCGGTGCTGGCCGCCGCCGGGAGAGC
CCGGGGCAGAGCAGAGGTGCTCATCAGCACTGTAGGCCCGGAAGATTGTGTGGTCC
CGTTCCCTGACCCGGCCTAAGGTCCCTGTCTTGCACTGGATAGCGGCAACTACCTC
TTCTCCACTAGTGCAATCTGCCGATATTTTTTTTTTGTATCTGGCTGGGAGCAAGA
TGACCTCACTAACCAGTGGCTGGAATGGGAAGCGACAGAGCTGCAGCCAGCTTTGT
CTGCTGCCCTGTACTATTTAGTGGTCCAAGGCAAGAAGGGGGAAGATGTTCTTGGT
TCAGTGCGGAGAGCCCTGACTCACATTGACCACAGCTTGAGTCGTCAGAACTGTCC
TTTCCCTGGCTGGGGAGACAGAATCTCTAGCCGACATTGTTTTGTGGGGAGCCCTAT
ACCCATTACTGCAAGATCCCGCCTACCTCCCTGAGGAGCTGAGTGCCCTGCACAGC
TGGTTCCAGACACTGAGTACCCAGGAACCATGTCAGCGAGCTGCAGAGACTGTACT
GAAACAGCAAGGTGTCCTGGCTCTCCGGCCTTACCTCCAAAAGCAGCCCCAGCCCA
GCCCCGCTGAGGGAAGGGCTGTCACCAATGAGCCTGAGGAGGAGGAGCTGGCTACC
CTATCTGAGGAGGAGATTGCTATGGCTGTTACTGCTTGGGAGAANGGCCTAGAAAG
TTTTGCCCCCGCTGCGGCCCCAGCANAATCCAGTGTTGCCTGTGGCTGGAGAAAGG
AATGTGCTCATCACCAGTGCCCTCCNTTACGTCAACAATGTCCCCCACCTTGGGAA
CATCATTGGTTGTGTGCTCAGTGCCCGATGTCTT

~~Sequence ID 1368~~SEQ ID NO: 463

CAGTGAGCCAAGATCACACCACTGCACTCCAGCCTGGACAACAGAACGAGACTCCA
TATCAAAAAAATTAAATTAAAATATAATAAATTTCTTGCCGGGCGCAGTGGCTCAC
ACCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAAGTCAGGAGA

Marked-Up Copy

TTGAGACCATCCTGGCTAATACAGTGAAACCCCGTCTCTACTATAAATACAAAAA
TTAGCTGGGCATGGTGGCGGGCGTCTGTAGTCCCAGCTACTCAGGAGTCTGAGGCA
GGAGAATGGTGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGTGCCACT
GCAATCCAGCCTGGGCAGCAGAACGAGACTCCATCTCAAATAAATAAATAAATAA
ATGAATTTTCAGCTAGAAGAGCCTTATTCCATTTTCCTTTTTATTAAACATCTGGCA
TAAGTTGGTAAGTATGTGAAGTTTATCATATATTCTTATGCGAATTATTATTTTCG
CCTTTTTTTTTTATAATTCTGTCTGGGATTTGAATAGTAGAGTTTGAATTCAGGAAG
GACACCTGTGATAGGACAATAAAAT

~~Sequence ID 1369~~SEQ ID NO: 464

CTGATTGCAAAAACATTACAACCTCAGTACTGCGGCTTTCATTCAAATAGGTTATAT
GTATAAACTGAGGTTCAACAATATTGTATTTGAGATGGGAAAGTTAAAGAAATGCA
ATAATGTAAATAATACTTAAGAAAATAAGATCTCAGGAACTGTGTATACTCTGTA
CTTTTATGCAACTTTTATCAGATCATTTTCAGTATATGCATCAAGGATATAGTGTATA
TGACATGAACTTTGAGTGCAAAAACGTACTATGTACCTTTTGTTTATTTTGCTGT
CAACATCTAAATAAAGGTTTTTTTG

~~Sequence ID 1370~~SEQ ID NO: 465

CGAAAGGACTACAGAGCCCCGAATTAATACCAATAGAAGGGCAATGCTTTTAGATT
AAAATGAAGGTGACTTAAACAGCTTAAAGTTTAGTTTAAAAGTTGTAGGTGATTAA
AATAATTTGAAGGCGATCTTTTAAAAAGAGATTAAACCGAAGGTGATTAAAAGACC
TTGAAATCCATGACGCAGGGAGAATTGCGTCATTTAAAGCCTAGTTAACGCATTTA
CTAAACGCAGACGAAAATGGAAAGATTAAATTGGGAGTGGTAGGATGAAACAATTTG
GAGAAGATAGAAGTTTGAAGTGGAAAACGTGAAGACAGAAGTACGGGAAGGCGAAG
AAAAGAATAGATAAGATAGGGAAATTAGAAGATAAAAACATACTTTTAGAAGAAAA
AAGATAAATTTAAACCTGAAAAGTAGGAAG

~~Sequence ID 1371~~SEQ ID NO: 466

GTCCAGNAGAAAGTTCAGTGACTTGTCCAGAGCTGCAGGTCTTAAGAGGCTGAAAT
CTCGCCTCTGCCTCGAGGCTGCGGTTCCACTGACCCATACTACTTGCCTTCAGGAA
AGAGAAATGGTGTAGGAAGGCTGTGGATGAAGACGCTTACATTCATGAAGGATTTG
GATAGGCGAACATGAGCTTTTCCACCAAATTTTCAGAATTTTAAGAAATGCCTTAAA
TTATTTCTTAAAAATCAATTTGGGGCAGACGAGAAGTTCTGATAATAGTTTTTAGG
GAACATGATAAAATTCTGACCTTAGAAGTGGTATACCAGTTTGAGAAGAAGAACAA
GCTATAAACGGTGTAGATAACATTCACGGCTATTTAAGAAAGAGTTACTAAGGGAA
ACCAGAATGACTTAAGAGTGTTACTCTTCTTTTTCTGAGAGAACAATAGCATCATC

Marked-Up Copy

TCAGAAAGCCTTTCATGCCATTAATAGGTAAGAATCTGGGCTTCTTGACCATGGG
TTAGACTTTCTTACAAAACCATAATATGCATTTCTAGCAAAATTTATGCTATTAC
ATTTCTTATCTCAACAAAGACTGGTAAATTCAGTACTTATTCCTCAATTTTCCTA
CCCTTAAATGGGGATATTCTGCCTCTCCAAGGAATGCTGGGAACAAGCAAGTCCT
CATGTTAGGGGTCTTTGAGTTTTTCATGGAAGTTTAGGTTATTTATATGATGACATA
GTTGTCAACTTACTTTCAGGATGGACTTTTCTTTTGTGAGTTTGTGACCTAAATAC
AATAGTTGTTATGCATGTCCAGTTTATGGAAGTACCACTGCAATANCAG

~~Sequence ID 1372~~SEQ ID NO: 467

CAGTGCAGCCAAGTATCACACCACTGCACTCCAGTCCTGGACAACAGAAACGANTA
CTCCATATCAAAAAAATTAAATTAAANGATAATAAATTTCTTGCCGGGCGCAGTGG
CTCACACCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAAGTCA
GGAGATTGAGACCATCCTGGCTAATACAGTGAAATCCCCGTCTCTACTATAAATAC
AAAAAATTAGCTGGGCATGGTGGCGGGCGTCTGTAGTCCCAGCTACTCAGGAGTCT
GAGGCAGGAGAATGGTGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGT
GCCACTGCAATCCAGCCTGGGCAGCAGAACGAGACTCCATCTCAAATAAATAAATA
AATAAAATGAATTTTCAGCTAGAAGAGCCTTATTCCATTTTCCTTTTTTATTAAACAT
CTGGCATAAGTTGGTAAGTATGTGAAGTTTATCATATATTCTTATGCGAATTATTA
TTTTCGCCTTTTTTTTTTATAATTCTGTCTGGGATTTGAATAGTAGAGTTTGAATTC
AGGAAGGACACCTGTGATAGGACAATAAAATCTA

~~Sequence ID 1374~~SEQ ID NO: 468

GAAAGCACATATGATATACATGTGTGTCATATGTATTATTTTGTGGCCATCTGAG
TCTTCAAATTTGTTACAGAATACCTGCATATTAATATTTCAAGGTATGGATTAAT

~~Sequence ID 1378~~SEQ ID NO: 469

CTGAGTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAA

~~Sequence ID 1380~~SEQ ID NO: 470

CCAAACCCAACCTGGTCCAGTAGGATACTCACCTTACAGGGGGCGTCTCAAGAGTCT
CACAGTTCCCTTGCGTCTTAAGAGACTCACTGTTGGACCAGGCGTGGTGACTCACG
CCTGTAAAACCAGCACTTTGGGAGGCCGAGGCGGGCGGATCAGTTGAGGTCAAGAG
TTCAAGACCAGCCTGACCAAGGTGCTGAAACCCCGTCTCTACTAAAAATACAAAAA
TTAGCCAGGCATGGTGGTGTGCGCCTGTAATCCCAGCTACTCCAGAGGCTGAGGCA
GGAGAATCTCTTGAACCCAGGAGGTGGAGGTTGCAGTGAGTCGAGATCATGCCACT
GCACTCCAGCCTGGGTGACAGAGCGAGACTCCGTCTTAGAAAAAAAAAAAAAAAAAA

Marked-Up Copy

AAAAGAACCTCACAGTTCAGCAGGGTTCTAGCATGAGACAATGAGGACAAGGGTAG
GTGAGCAGGTGGAAAGAGTGAGAACAGGTCAATTGTGATGGAGAAAATAATAAAGA
CAGAAAAGGCAGAAGACTGCCTGGCAGAAGACCTGTCCCAGCAGATACAAAAATAC
AGACAACAGGAGCCAGCATAGACCCCTTGACCTGTGTAAGTCTTTCTCAGGCCTTCT
TTTAAGTAGAAACATGCCTTTGAAAAAAGTTTTTAATAAACAGGAAAATCATAAAT
CCCTATTTACATAAATAATATATCCTGGTCTTATTCTTAAAACCATTGATTTTTTCA
CGGCTCATTAANAAAGCTGGGCGAGGTGGCTCACGCCCGTCATCCTAGCACTTTGG
GAGGCCGAGGCGGGCANATCACAAGGTGAGGAGTTGGGAGACCAGCCTGACCAACA
CGGTGAAACCCAGTCTCTACTAAAAATACAAAAATTANCTGGGGGTGGTGGTGTGT
GCCTGTAATCCAAGCTACTCGGGAGGCTGAGGCAGGA

~~Sequence ID 1382~~SEQ ID NO: 471

CTTACTACCTCCAACATGAAACAAGCAGCCCCGCACTTCTCGAAGGTCTGAGTTAC
TTGGAATCGTTTTTACCACATGATGGACAGAAGGAATATTTTCAGATATCTCTGAAAA
CCTCAAGCGTTACCTTCTTCAGTATTTTAAGCCAGTGATTGACAGGCAAAGCTGGA
GTGACAAGGGCTCAGTCTGGGACAGGATGCTCCGCTCGGCTCTCTTGAAGCTGGCC
TGTGACCTGAACCATGCTCCTTGCATCCAGAAAGCTGCTGAACTCTTCTCCCAGTG
GATGGAATCCAGTGGAATAATAATATACCAACAGATGTTTTAAAGATTGTGTATT
CTGTGGGTGCTCAGACAACAGCAGGATGGAATTACCTTTTAGAGCAATATGAACTG
TCAATGTCAAGTGCTGAACAAAACAAAATTCTGTATGCTTTGTCAACGAGCAAGCA
TCAGGAAAAGTTACTGAAGTTAATTGAACTAGGAATGGAAGGAAAGGTTATCAAGA
CACAGAACTTGGCAGCTCTCCTTCATGCGATTGCCAGACGTCCAAAGGGGCAGCAA
CTAGCATGGGATTTTGTAAAGAGAAAATTGGACCCATCTTCTGAAAAAATTTGACTT
GGGCTCATATGACATAAGGATGATCATCTCTGGCACAACAGCTCACTTTTCTTCCA
AGGATAAGTTGCAAGAGGTGAACTATTTTTTTGAATCTCTTGAGGCTCAAGGATCA
CATCTGGATATTTTTCAAACGTCTCTGGAAACGATAACCAAAAATATAAAATGGCT
GGAGAAGAATCTTCCGACTCTGAGGACTTGGCTAATGGTTAATACTTAAATGGTCA
ATAGAAAAAGTAGGCTGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGA

~~Sequence ID 1387~~SEQ ID NO: 472

AAAATT
TTTCAGT
GTTAAAGTAGGTTTGTGACGCGGCCACGAATTTCCCGGGGACCAA

~~Sequence ID 1389~~SEQ ID NO: 473

TTTTTTTTTTTTTTTTTGGGAGTCAGTTTTCTTTCTTTCTTTCTTTCTTTTTTTTTTTT

Marked-Up Copy

GNTTTTCGGAACGGAGTCTCGCTTTCTCGCCCACTCTGGAGTGGNGCAGTGGGGN
GGTCTCAGCTCACCACAGCCTCCACCTCCTGGGCCCCAAGCGATCCTNTCACCTCAG
CCTCCTGCGTAGCTGGGACTACAGGCGTGCACCACCATTCCCAGGTAATTTTTTGTA
TTTTTTGTANANACAGGGTTTCACTGTTGTTGCCCAGGCTGGTCTCGAACTCCTGC
TTCAGTCTGCCANAATGCTGGATTCTAGGCGTGAGCCACCGNGCCTGGCCCCAAAAG
TTACTTTTCTTACAGAAGCAAAGCTTTAATGCATTTTACTGAATGCTTATAGCTTT
GTAGATACTGAAAAGAGTATGAGCGTCACATACAGACACATNTAACAGCACTGCCT
CCAACCAGCCCCCTACCCACTGGTCAGGNGAGTAANAATCAAAATTCTTTTCTGNGA
GTGGAACGGAAATTTTCTCTCTCCTCCTCAGGCAAGTAGTTAANAGGCTGGNGGGA
GTCATGGCCCCATTTTGTTCAAAATACAAGCTCCACAGGAACAAAAGGCTGAACTG
CTCACCTCCCAACTGATGAACCTCGTCTTTGTTCCATGTCAAAGGGGCCCTTTGTGT
TACTGCAGCAGAACTCCAGCTATCAAACCATCAGGCACCAAAAGTAAACTCCTT
TCTCTAAAAAGACCTCTCTTTACCTGAGCCTTTCAATGCATCTTTGCCCCCANATA
ATCCTGGATGAGATAATCCCCAGAGGAANACCAGCGCTTGCCTAGTGAAATTATAC
TATGAGACAAGGGTAAAAGACCTCAANACCGGGTTGGCAGGTAAGGGAGTAGGGN

~~Sequence ID 1390~~SEQ ID NO: 474

TCNGTGGCACCCGTTTCCGGCACCTTCAGACTCTGAAGAGCCACCTGCGAATCCAC
ACAGGAGAGAAACCTTACCATGTACGTAAGCCTCTTGAGGCCGCTCTCTGACCTGC
GGGGATGTGGAGGGCAGGGAAGGAGGTGGAGCGCAGGGAAGGAGGTGGAGCAGGGA
GGCAGTGGAAGTGTGCTCCCATCTCAAGCACACAGTGGGGCAACCACTACGCTA
ATGGTTGGAAGACCTAGATCTGGGCCCAATGGCCAGACACCCTGCTTGACCTTGGC
CCAAGCATTAGGGGACTCATCTTTAAAATGAGGGTATGGGACTAGATGATCTGGGC
CTTAGGAGAGGAGT

~~Sequence ID 1391~~SEQ ID NO: 475

CGGCTNCTACCCTGCGGAGATCACACTGACCTGGCAGTGGGATGGGGAGGACCAAA
CTCAGGACACCGAGCTTGTGGAGACCAGGCCAGCAGGAGATGGAACCTTCCAGAAG
TGGGCAGCTGTGGTGGTGCCTTCTGGAGAAGAGCAGAGATACACGTGCCATGTTCA
GCACGAGGGGCTGCCGAGCCCCCTCACCTGAGATGGAAGCCGTCTTCCCAGCCCA
CCATCCCCATCGTGGGCATCGTTGCTGGCCTGGCTGTCCTGGCTGTCCTAGCTGTC
CTAGGAGCTATGGTGGCTGTTGTGATGTGTAGGAGGAAGAGCTCAGGTGGAAAAGG
AGGGAGCTGCTCTCAGGCTGCGTCCAGCAACAGTGCCCAGGGCTCTGATGAGTCTC
TCATCGCTTGTAAGCCTGAGACAGCTGCCTGTGTGGGACTGAGATGCAGGATTTT
TTCACACCTCTCCTTTGTGACTTCAAGAGCCTCTGGCATCTCTTTCTGCAAAGGCA
TCTGAATGTGTCTGCGTTCCTGTTAGCATAATGTGAGGAGGTGGAGAGACAGCCCA

Marked-Up Copy

CCCCCGTGTCACCGTGACCCCTGTCCCCACACTGACCTGTGTTCCCTCCCCGATC
ATCTTTCCTGTTCCAGAGAAGTGGGCTGGATGTCTCCATCTCTGTCTCAACTTCAT
GGTGCGCTGAGCTGCAACTTCTTACTTCCCTAATGAAGTTAAGAACCTGAATATAA
ATTTGTTTTCTCAAATATTTGCTATGAAGGGTTGATGGATTAATTAAATAAGTCAA
TTCCTGGAAGTTGAGAGAGCAAATAAAGACCTGAGAACCTTCCANAATCCG

~~Sequence ID 1392~~SEQ ID NO: 476

TGAAACAAAATGAATTTNTATGGGTAAGAGAGGGTAATATTTTAGAGTTGTGTTAC
AAAAC TACAAATTTTTATTAAATTAATAAATCAGAATACTAAATCCATGTGTTTTT
TTCTTTCTTAAAAAATATCTTTTGGCTGGGCACGGTAGCTCATGGCTGTAATCCCA
GCACTTTGGGAGGCTGAGGTGGGTGGATCGCCTGATGTCAGGAGTTCAAGACCAGC
CTGGTCAACATGTTGAAACCCCATCTCTACTAAAAATATAAAAAATTAGCCGGTGTG
GTGGTGGGCGCCTGTAATCCCAGCTACTCAGGAGGCTAAGGCAGGAGAATTGCGTG
AACCAGGAGTTCAGTGATGTAGCGGGGAGCTGAGATTGTGCCACTACACTCCAGC
CTGGATGACAGAGTGAGACTCCATCTCAAAAAAAAAAAAAAAAAAAAAA

~~Sequence ID 1394~~SEQ ID NO: 477

GCATAATGTGAGGAGGTGGAGAGACAGCCCACCCCCGTGTCCACCGTGACCCCTGT
TCCCATGCTGACTTGTGTTTCCTCCCCAGTCATCTTTCCTGTTCCAGAGAGGTGGG
GCTGGATGTCTCCATCTCTGTCTCAACTTTATGTGCACTGAGCTGCAACTTCTTAC
TTCCCTACTGAAAATAAGAATCTGAATATAAATTTGTTTTCTCAAATATTTGCTAT
GAGAGGTTGATGGATTAATTAAATAAGTCAATTCCTGGAATTTGAGAGAGCAAATA
AAGACCTGAGAACCTTCCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

~~Sequence ID 1395~~SEQ ID NO: 478

CTTACCATGTCAGTGCACAGAAATGCTGTCTTGGGATGTAGGAAAAATAAATCCAC
AAAAGCTACCAAGTTTGAAGGGGACCATGAGTCTTCAGGCTGGAGCTTCCAAACCA
GATGAAAACCCCACAATTAACCTGCAGTTTAAGATCCAGCAGCTGGCCATTTCTGG
ACTCAAGGTGAATCGTCTGGATATGTATGGAGAAAAGTACAAACCTTTAAGGGCA
TAAAATACATGACCAAAGCTGGGAAGTTCCAAGTTCGAACCTGAAGGGAGCATTTG
CTGAGGGAATAGTCTTGACATTTTTTTCATTTCTTACTTGTCTAAAAGTAAAAAA
AATATCAGCCTGTCTCCTAGGTCAGTCCCCTCCTGGACCCACCCGCTCCCTTTTTT
CCTTAGCCTTCAGTGCCATGGAAC TAATCAAGGGAGGAAAAGGTCACCAGGGAGAA
CTGGACAGAACTGAAACACAGCAACACCAGTTCTCAAGGACAAGGTGTGTGATGGG
GGTAGGAAGCTTGGTGCTTATGTAACCATTTTAAACGTGGTTTCTATAGGAAAGAC

Marked-Up Copy

CAACATTTGTTTAGCTTGCTTGGCTTTAATTATCTAAAGCCAATGAAAGACTTCTT
TGTTGATTTTTTAAGATAGAAAGATT

~~Sequence ID 1396~~SEQ ID NO: 479

CAAACACTATGTTATTTTATGAANAAGACTTGAACATCTATGGATTTTGGTATTTG
CAAGGGGTGAATGGGGTATTTGCAAGCAGTGAATGAGGAGGCCTGGAACCAATCTT
CTGCTGATATTGAGGCACAACCTGAAAAAGGTATATTACTTAAATCTCTTATTGTAT
TGTAACCTGTATAAGTAATGAAATTAAAAGGCAGAAATTGTCAGACTGAATAAAAT
GAAAAGACCAAACAATATGCTGCTTACAAGAAACACAATTCAAATATAAGGACACA
ATTAGTTTAAAGGAAAAGAACTGGAAAAGATATACCATGATAACACAAGTCAGAAG
AAAGCTGCTGTGGATATATTAATATGAGATGTAGATTTTCAGAGCAGTGAATATTGC
CAGGCATAAAGAAAGTTATTACATAATAATTAAGGTATCAGTTCATCAAGAAGATG
TAATAACCCTAAGTATTTATACAACCTAATATCAGAGCTTCAAATACATGAAGCAA
AAACCAGTGGAATTGATAGGAGAAACACACAATTACACAATTATAGTCAGAATTTT
CAACATATCTTTCTCAATGGAGAAAACAACTAGACAGGAAATCATTAAAGGATATAG
ATGATTTTAAATTATATGATCAACTACCTGGACGTAATTGGCATTATGGAACACTG
CACCACCAACAGCAGAGTACATATTATTTTCAAGTACACAGAAAACAGTTACCAAT
ATAGACCATTTTCTGGGTCATAAAACACATCTCAATAAATGTAAACAATTAATGT
TATATAAAGTATGTGCTCTGACCNCAAAGGAATTAGAGATCAATAAAAGAACATCT
TTGAAAAATCTCACNTATTTAAAACTAATAACTCACTTCTAAATAACTCCTGTNT
CAAGAGAATNAAANGG

~~Sequence ID 1397~~SEQ ID NO: 480

CCCAGCCTCACTGCGCCCCGTCAGGCCAGGCAGCTGCCCTCAGGGTCTGCCAAGGT
GGGGGTCAAGGGCCATGGGGGCAGGTAGCTCTGCCTGCAAAGCCCACAAGCATGTC
AGATCACCTGGGCTGCAGACAGACAAACACCTGAGCTGTTCTGAATACCTTCAGGT
TCCTGGCCTCGCTGAGCAAGTGCAGAAATTTTACCTTCAAGGATCAGGGTTTTTC
TGTTTGTTTGTTTTTTTAACACACACATATGTGAACAAAGAGTATGCGTTTGTACTG
GCAGAAGAAGCGTCTGGTAAGACAACCAGCAAGTTAACAATGGTCACCTCCAGAAA
TGGGCTGGGTAAACCAAAGAATTTTTTTGTTTTTTTGTAGTCAGGGTCTA
GCTCTGTCACCCAGGCTGGAACGCACTGGTGTGATCACGGCTCACTGCAGCCTTGA
CCTCCCTGGCTCAAGCAATCCTCCCAGCTCAGCCTCCTGAGTCGTTGGGACTACAG
GCACGTGCCACCACGCCTGACACATTTTTTAAATTTTTGTAGAGACAGTGTTTCAC
CATGTTGCCCAGGCAGGTCTCAAACCTCTGGGCTCAAGTGGTCCTCCAGCTTCAGC
CTCCCAAAGTGCTAGGATTATAGGTGTGAGCCACAGTGCCCAGCCCCGTAGTGGAG

Marked-Up Copy

AATTTCTGTTGAATGAACCAAAAGCAACTGCCAACCTCTCCATGCACCATGTGTTT
CAGAGGAGAAAGCACAGTGAAGAATGCAGTGTGTTCTGAGGTCCTGTCACCCCTGA
GGCTGTGTGTGTCCTTTGCCAAATTAAAGAGTCTTACTGAATGCGGTGCATCCAGG
AGACAGGCCNAGGTTTGGACTGGTAAAAAAAAA

~~Sequence ID 1399~~SEQ ID NO: 481

CAGACACCTGGNAGAACGGGAAGGAGACGCTGCAGCGCGCGGACCCCCCAAAGACA
CATGTGACCCACCACCCCATCTNTGACCATGAGGCCACCCTGAGGTGCTGGGCCCT
GGGCTTCTACCCCTGCGGAGATCACACTGACCTGGCAGCGGGATGGCGAGGACCAAA
CTCAGGACACCGAGCTTGTGGAGACCAGACCAGCAGGAGACAGAACCTTCCAGAAG
TGGGCAGCTGTGGTGGTGCCTTCTGGAGAAGAGCAGAGATACACATGCCATGTACA
GCATGAGGGGCTGCCGAAGCCCCCTCACCCCTGAGATGGGAGCCATCTTCCCAGTCCA
CCGTCCCCATCGTGGGCATTGTTGCTGGCCTGGCTGTCCTAGCAGTTGTGGTCATC
GGAGCTGTGGTCGCTGCTGTGATGTGTAGGAGGAAGAGTTCAGGTGGAAAAGGAGG
GAGCTACTCTCAGGCTGCGTCCAGCGACAGTGCCCAGGGCTCTGATGTGTCTCTCA
CAGCTTGAAAAGCCTGAGACAGCTGTNNTGTGAGGGACTGAGATGCAGGATTTCTT
CACGCCTCCCCTTTGTGACTTCAAGAGCCTCTGGCATCTCTTTCTGCAAAGGCACC
TGAATGTGTCTGCGTCCTTGTTAGCATAATGTGAGGAGGTGGAGAGACAGCCCACC
CTTGTGTCAACTGTGACCCCCTGTTCCCATGCTGACCTGTGTTTCCCTCCCCAGTCA
TCTTTTTTGTTCNCAATAGGTGGGGCCTGGATGTCTCCATCTCTGTNTCA

~~Sequence ID 1440~~SEQ ID NO: 482

TTATAAGGTACTTTTAAGGTATTTTAGTTGTCTTAGTCTATATTTCTGTACTCACC
TTTCTTTATCCACTCATCAGTTGATGGGCATGTAGGTTGGTTCCATATCTTTGCAA
TTCTGAATTGTGCTGTGATCAGGTGTCTTTTTAGTATAATGATTTACTCTCCTTTG
GGTAGATACCCAGTAGTGGGATTGCTGGATCGAATGGTTTTTATAATTTTCTATTT
TACCACAGTTTCTCTCTGCATTTTTCTCTTTGACCACTAACCATGTGAAATTCTC
ATATTGACCTTTATAATGATCATGAACTCTTAGTATCATTGGGAAGGCCACATTTG
CCACTTATGATTGTAAACCTTATCCTCCATTTTTCTCTGTTATTGTTGGTGCAAAAA
GCACCTATTATACCAGGACTTTAAAAATCAGTCTGATAAGTCTTTGATAAGTCTAA
TAATAATAACTGATAAGTCCATTGAATTTGCTTCTGATTACTTTTTCTTTAGTAGC
TAAACATGTATGTACTCCTATGATTACAATGAACACTCCTCTCCATTTAAATTAAT
TATTTACATTGATGAAATAGCAAAATGTTAATGACTAAATACTGTCTTGGTTTTTT
CGTTCCAGGTCAGTCAATATTAACCTTCTTATAATTTTCTTTTTTTTTCTTT

~~Sequence ID 1447~~SEQ ID NO: 483

GCAAGGACTAACCCCTATACCTTCTGCATAATGAATTAAGTAGAAATAACTTTTGCA
AGGAGAGCCAAAGCTAAGACCCCCGAAACCAGACGAGCTACCTAAGAACAGCTAAA
AGAGCACACCCGTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGGCGACA
AACCTACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAACTTT
AAATTTGCCCACAGAACCCTCTAAATCCCCTTGTAATTTAACTGTTAGTCCAAAG
AGGAACAGCTCTTTGGACACTAGGAAAAAACCTTGTAGAGAGAGTAAAAAATTTAA
CACCCATAGTAGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACAC
CCACTACCTAAAAAATCCCAAACATATAACTGAACTCCTCACACCCAATTGGACCA
ATCTATCACCCCTATAGAAGAACTAATGTTAGTATAAGTAACATGAAAACATTCTCC
TCCGCATAAGCCTGCGTCAGATTAAACACTGAACTGACAATTAACAGCCCAATAT
CTACAATCAACCAACAAGTCATTATTACCCTCACTGTCAACCCAACACAGGCATGC
TCATAAGGAAAGGT

~~Sequence ID 1448~~SEQ ID NO: 484

GGCCACCGGGTGCAAGGTCAGGGCTGGGGTGGAGGCTGGGAAGCCCAGGGCTTGGC
CCACTGTGGCCGCCTTGTGTGGTCACTGCTTTCCTGGGCCTGCTGTGAGCTCCCTC
TAGGACCCCAGGCCTGTCTGGTGGGTCACTGTGACCACCACCTTGCACAGCACCTG
GCGCGTGGCAGGTGCTCAAACATTACTTGTTCGGAATGAACTTCATCTTGCTCTT
GGCTTTTTGACTAATGCTGTGGAACATCTGACTAATTAGTGACTCTTTGGGGCCCC
CAGTTTCCCAGCTATAAAGTGGTAATATTAAGATAATAATTCGGCCGGGCGCGGTG
GCTCACGCCTGTAATCCCAGCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACGAG
GTCAGAAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCATCTCTACTAAAAA
TACAAAAAATTANCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCANGAG
GCTGANGCAGGAGAATGGTGTGAACCCGGGAGGCAGAGGTTGCAGTGAACCAAGAT
CGNNCCACTGCACTCCAGCCTGGGCAACAGAGCGAGACTCCATCTTAAAAAA

~~Sequence ID 1449~~SEQ ID NO: 485

AATCAGGGCCGCAGTGTGTTCTGCGCCTGCCAGAGCTGACTCCTGATTTAACCGC
TGGCGTAACCGCGGGTTGCACGCATGCGTGCTGAAAAGCCTTTCACCCTCACGTGG
TTTCTTTTTTAACCAGTCATCAAGCGAGGCTCGCGCGCAGGCCCGCGTTGGAAAA
TGGCGGGGAAGCTGAAACCTCTGAATGTGGAGGCGCCAGAAGCTGCTGAGGAGGCT
GAAGGTAGTGAGGGCAAGTGGGCTGCACTCCTTTCTCTCCAACCAGGGCAGAAAGG
AGGGAGGATTCGTCCCATTACAATAATGAAATAATGATATTCTAATTTTTTTTAAAT
AAAATGTTAAGCCTTTTGTATTGAA

~~Sequence ID 1450~~SEQ ID NO: 486

GGAAANCATGAGGCTTCGGGAGCCGCTCCTGAGCGGCAGCGCCGCGATGCCAGGCG
CGTCCCTACAGCGGGCCTGCCGCCTGCTCGTGGCCGTCTGCGCTCTGCACCTTGGC
GTCACCCTCGTTTACTACCTGGCTGGCCGCGACCTGAGCCGCCTGCCCCAACTGGT
CGGAGTCTCCACACCGCTGCAGGGCGGCTCGAACAGTGCCGCCGCCATCGGGCAGT
CCTCCGGGGAGCTCCGGACCGGAGGGGCGCCGCCGCCCTCCTNTAGGCGCCTCC
TCCCAGCCGCGCCCCGGGTGGCGACTCCAGCCCAGTCGTGGATTCTGGCCCTGGCCC
CGCTAGCAACTTGACCTCGGTCCCACTGCCCCACACCACCGCACTGTCGCTGCCCG
CCTGCCCTGAGGAGTCCCCGCTGCTTGGTAAGGACTCGGGTCGGCGCCAGTCGGAG
GATTGGGACCCCCCGGATTTCCCCGACAGGGTCCCCCANACATTCCCTCAGGCTG
GCTCTTCTACGACAGCCAGCCTCCCTCTTCTGGATCAGAGTTTTAAATCCCANACA
GAGGCTTGGGACTGGATGGGAGAGAAGGTTTGCAGAGTGGGTCCCTGGGGAGTCCT
GTTGGAGGCGTGGGGCCGGGACCGCACAGGGAAGTCCCGAGGCCCTCTAGCCCCA
AAACCANAGAAGGCCTTGGAGACTTCCCTGCTGTGGCCCGAGGCTNAGGAAGTTTT
GGAGTTTTGGGTCTGCTTANGGCTTCNAGCAGCCTTGCACTGAGAACTTTGGTAGG
GACCTCGAGTAATCCACTCCNTTTTNGGGACTGACGTGAGGCTCCCGGTGGGGAAA
GANACTGACCTNTC

~~Sequence ID 1453~~SEQ ID NO: 487

CCGACCTGTCTCGCTCCGTGGCCTTAGCTGTGCTCGCGCTACTCTCTCTTTCTGGC
CTGGAGGCTATCCAGCGTACTCCAAAGATTAGGTTTACTCACGTCATCCAGCAGA
GAATGGAAAGTCAAATTTCTGAATTGCTATGTGTCTGGGTTTCATCCATCCGACA
TTGAAGTTGACTTACTGAAGAATGGAGAGAGAATTGAAAAAGTGGAGCATTCAGAC
TTGTCTTTCAGCAAGGACTGGTCTTTCTATCTCTTGTACTACACTGAATTCACCCC
CACTGAAAAAGATGAGTATGCCTGCCGTGTGAACCATGTGACTTTGTACAGCCCA
AGATAGTTAAGTGGGATCGAGACATGTAAGCAGCATCATGGAGGTTTGAAGATGCC
GCATTTGGATTGGATGAATTCCAAATTTCTGCTTGCTTGCTTTTTAATATTGATATG
CTTATACACTTACACTTTATGCACAAAATGTAGGGTTATAATAATGTTAACATGGA
CATGATCTTCTTTATAATTCTACTTTGAGTGCTGTCTCCATGTTTGATGTATCTGA
GCAGGTTGCTCCACAGGTAGCTCTAGGAGGGCTGGCAACTTAGAGGTGGGGAGCAG
AGAATTCTCTTATCCAACATCAACATCTTGGTCAGATTTGAACTCTTCAATCTCTT
GCACTCAAAGCTTGTTAAGATAGTTAAGCGTGCATAAGTTAACTTCCAATTTACAT
ACTCTGCTTAGAATTTGGGGGAAAATTTAGAAATATAATTGACAGGATTATTGGAA
ATTTGTTATAATGAATGAAACATTTTTGTGATATAAGATTCATATTTACTTCTTAT
ACA

~~Sequence ID 1454~~SEQ ID NO: 488

TAAATAGGGAATCCTTTCCCCATTGCTTGTTTTTCTCAGGTTTGTCAAAGATCAGA
TAGTTGTAGATATGCGACGTTATTTCTGAGGGCTCTGTTCTGTTCCATTGATCTAT
ATCTCTGTCACATGCACACGTATGTTTGTTGTGGCACTATTCACAGTGGCAAAGAC
TTGGAACCAACCCAAATGTCCAACAATGATAGACCGGGTTAAGAAAATGCGGCACA
TATACACCATGGAATACTATGTAGCCATAAAAAATGATGAGTTCGTGTCCTTTGTA
GGGACATGGATGAAATTGGAAATCATCATTCTCAGTAACTATCGCAGGAACAAAA
AACCAAACACTGCATATTCTCACTCATAGGTGGGAATTGAACAGTGGGAACACATG
GACACAGGAAGGGGAACATCACACTCTGAGGACTGTTGTGGGGTGGGGGGAGGGAG
GAGGGATAGCATTGGGAGATATACCTAGTGCTGGATGACGAGTTAGTGGGTGCAGC
GCACCAGCATGTCACATGTATACATATGTAACCTAACCTGCACATTGTGCACATGTA
CCCTAAAACTTAAGGTAT

~~Sequence ID 1456~~SEQ ID NO: 489

CCGCAACAAACACGGGAGTGCAGATATCGCTGCGATGGGCTGATTTTCCTTTATTTG
GGTATATACCCAGCAGTGGGATTGCTGGATTGTATGGTAGCTCTATTAGTTTTTTG
AGGAACCTCCAAACTGTTCTNCATAGTGGTTGTACTCATTTACATTCCCAGTGTGA
ACCCTGAAAATTTGAGGCAGGTCTCAGTTAAATTAGAAAGTTGATTTTGCCAAGTT
GGGGACACGCACTCGTGACACAGCCTCAGGAGGAACTGATGACATGTGCCCAGGTG
GTCAGAGCACAGCTTGTTTTTATACATTTTAGGGAAACCTGAGCCATCAATCAACA
TACGTAAAATGGGCCGGGCACAGCAGCTCAAGCTGTAATCCCAGCACTCTGGGAGG
CCGAGGCGGGTGGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGG
TGAAACCCCGTCTCTATTAAAAATACAAAGCTTAGCTGGATGTGGTGGCGCATGCC
TGTAGTCCCAGCTGCTCTAGGAGGCTGAGGCATGAGAATTGCTTGAACCTGGGAGG
CAGAGGCTGCAGTGAGCCGAGATCGAGCCACTATACTCCAGCCTGGTCAACAGAGT
GAGACCCTGTCT

~~Sequence ID 1460~~SEQ ID NO: 490

CCACAACCTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACCTGACTCCTG
AGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGT
GGTGAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGA
GTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGG
CTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAAC
CTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGA
TCCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACT

Marked-Up Copy

TTGGCAAAGAATTCACCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGT
GTGGCTAATGCCCTGGCCCACAAGTATCACTAAGCTCGCTTTCTTGCTGTCCAATT
TCTATTAAAGGTTCCCTTTGTTCCCTAAGTCCAACCTACTAACTGGGGGATATTATG
AAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTG

~~Sequence ID 1490~~SEQ ID NO: 491

ATGGGCATCTCTCGGGACAACCTGGCACAAGCGCCGCAAAACCGGGGGCAAGAGAAA
GCCCTACCACAAGAAGCGGAAGTATGAGTTGGGGCGCCCAGCTGCCAACACCAAGA
TTGGCCCCCGCCGCATCCACACAGTCCGTGTGCGGGGAGGTAACAAGAAATACCGT
GCCCTGAGGTTGGACGTGGGGAATTTCTCCTGGGGCTCANAGTGTTGTACTCGTAA
AACAAGGATCATCGATGTTGTCTACAATGCATCTAATAACGAGCTGGTTCGTACCA
AGACCCTGGTGAAGAATTGCATCGTGTCTCATCGACAGCACACCGTACCGACAGTGG
TACGAGTCCCCTATGCGCTGCCCCCTGGGCCGCAAGAAGGGAGCCAAGCTGACTCC
TGAGGAAGAAGAGATTTTAAACAAAAACGATCTAAAAAAATTCAGAAGAAATATG
ATGAAAGGAAAAAGAATGCCAAAATCAGCAGTCTCCTGGAGGAGCAGTTCCAGCAG
GGCAAGCTTCTTGCGTGCATCGCTTCAAGGCCGGGACAGTGTGGCCGAGCAGATGG
CTATGTGCTAGAGGGCAAAGAGTTGGAGTTCTATCTTAGGAAAATCAAGGCCCGCA
AAGGCAAATAAATCCTTGTTTTGTCTTCACCCATGTAATAAAGGTGTTTATTGTTT
TTGTT

~~Sequence ID 1491~~SEQ ID NO: 492

CTTNCACATACTGATTGATGTCTCATGTCTCTCTAAAATGTGTAAAACCAAGCTGT
GCCCCAACACCTTGGGNACATGTGGNGAGGACCTCCTGAGGCTGTGTCTATGGGCA
CACCTTAACCCTGGGAAAATAAACTTTCTAACTGACTTGAGAGCTGTCTCAGATA
TTCTGAGCTTACAGTTATTGTGAAATCATTTTTAATTATAAATTAAGTGGAGATTTA
CTTAAATCATGTGTAGTAGCCTGTGATATAGTCCTAGATACATACATTATCA
TCTTATGTATCTTCCCTCCCTCTTCCAGGTTCTGATAAAAACAGATGAAATCTGAA
AGACCATGACAGTAGTATTTTGAAAATGACAGTATTTGAAATTAAAAAATTGTAAA
AGTGTTCTGTTCTATCACTGCCAAAGGATAAGTTACAAATTGGTTCTTGGAACGTA
ATATGTACTATGTGCTTGCTATTTAATAATTTACCAGTCTTAGTCTTTTTTTATTCA
GACTAATTTTACCTTTTTTTTAACCTATGACTCTTTAGTTATAGTAGTACAAAAAAG
TAGTTTTAGTTATAGTTTTAGTTGTAGTACAAAAAGCATTTTCTGTAAGCTTAAT
TTCTTTCCCTTCCCGCTTTCCAGTCAGATGACTTTAGTGATTTGGAGTTGTGTG
CTTTATAAGTGCATTCCTCAGAGGACTTAATATTACTAAGATTTTAGCAACNCTGA
AATATGTT

~~Sequence ID 1492~~SEQ ID NO: 493

TGTNCCTGTAGTCCTGTGTGGGAGGATTGCCTGAGCCTAGGAGCTCAAAGTTGCAG
TGAGCCCAGATCGNGNCATTGCAGTCCAGCCTGGGTGACAGAGTGAGACCCCATGT
CAAAAAAAAAAAAAACAAAAACAGGGGGCCTGCCTCANCCAGCAGGTGAGGTCTGCC
ACTGAGAGCACTTCTAGCAGCAGGAACAGCCTCCACCCCCACACTGCAATCAAGTT
TTTTGGGTGAGCCTTAGGAGCTAANAAAGGGCCTAGTTTGNCTAAATAGCAGGAGT
TATATCCAGGGATCTTCAGGCCCAGGAATGCTAATGAGTAGGCATTCCATGGGCCC
TGGGAATGGCTTTGTGTGCCANAAATGATGGCCACAAAGGCCTTGCTGCCTTTTTT
CAAAATGGCTGCATCCAGCTGAGTGCTCTCTGCCAAAGGGGANAANAAAATAAGTC
TCCAGTGCATTTAGATTGGTCTCTCATCATCTCTCTCCTTTTTGTGTTTTATTAGTC
TCCTTAACCAAACTGCCAAGAAAGGCTTGGAATTGAAACAAAACCTGATANAANA
GGTAAGAGGTTGTTCTTTT

~~Sequence ID 1493~~SEQ ID NO: 494

TGTNTCAAAAAAAAAAAAAAGAACGGNAATGTACTGGAGATGTATTTGATAACCAA
GGNTTTAGGTAAATTTTCACCAGTATTAGTTNTATTTGCAAACCTGAAAAATGTTGT
AGGCTTAATATAAAATAACCACATTAGTGAACATTATATCTCTTAGAAGAAAGGCC
ATATTTTGCTCCTGCTTCTGTAAAAATATTATTTGTTTGAAGGGGAAATAATGGTA
GTGTGACCTTTCACTTAATTCCTACTCCCTTAATGTGAGAGAGACAAAATGAGCTG
AAGAAGGAAAATTCTGGAGTTACACTCCACAACCTTGAACATACTGACGGACATCT
CTGTTTTGACAACGATTTCTCCATGCCACCCATGCTNTAATGCCTTGTTGGATCACG
GACAACCCTCTTTGCACAAGCTACAGCATCAGCGATGTTATCTTGCAGCAAAGCAC
TGCAGGATAAATGACAGGCATTAACCTGCTCCTGGGGTTTTGCCATCATTACACCAG
TAGCGGCTATTGATCTGAAATATCCCATAATCAGTGCTTCTGTCTCCAGCATTGTA
GTTTGTAGCTCGTGTGTTGTAACCACTCTCCCATTTGGCCAAACACATCCAGTTTG
CTAGGCTGATTCCCCTGTAGCCATCCATTCCCAATCTTTTCAGAGTTCTGGCCAAC
TCACACCTTTCAAAGACCTTGCCCTGGACCGTAACAGAAAGGAGGACAAGCCCCAG
AACAATGAGAGCCTTCATGTTGAC

~~Sequence ID 1494~~SEQ ID NO: 495

TTGGTACCCGGGAAATTCTTTGCCGCGTCGACGGCCGGTGAGGCAGATCACCTGAG
CCCAGGAGTTCAGGACCAGCCTGGGCAGCATACCGGGATTCCATCTNNACTAAAAA
CAGTAGGCTGGGTGTGGTGGCTCATGTCTGTAAGCTCAGGACTTTGGAAGGCCAAG
ATGGGAGGATCACTTGAGCCTGGGAGTTTGACACCAGCTTGAGCATCGTAGCCAGG
CCCTGACTCTACAAAAAAGTGAAATAATTAGCCGAGTGTGGTGGTTCACACCTGTA
ATCCCAGCTGCTCAGGAGGCTGAGGTAGGAGAATCATTTGAACCCGGGAGGTGGAG

GTTGCAGTTAGCCGAGATCACGCCATTGCACTCCGGCCTGGGCGATAAAGCGAGAC
TCTGTCTCAAAAAAAAAAAAAA

~~Sequence ID 1495~~SEQ ID NO: 496

ATTCGGGCCGAGATGTCTCGCTCCGTGGCCTTAGCTGTGCTCGCGCTACTCTCTCT
TTCTGGCCTGGAGGCTATCCAGCGTACTCCAAAGATTCAGGTTTACTCACGTCATC
CAGCAGAGAATGGAAAGTCAAATTTCTGAATTGCTATGTGTCTGGGTTTCATCCA
TCCGACATTGAAGTTGACTTACTGAAGAATGGAGAGAGAATTGAAAAAGTGGAGCA
TTCAGACTTGTCTTTTCAAGCAAGGACTGGTCTTTCTATCTCTTGTACTACACTGAAT
TCACCCCCACTGAAAAAGATGAGTATGCCTGCCGTGTGAACCATGTGACTTTGTCA
CAGCCCAAGATAGTTAAGTGGGATCGAGACATGTAAGCAGCATCATGGAGGTTTGA
AGATGCCGCATTTGGATTGGATGAATTCCAAATCTGCTTGCTTGCTTTTTAATAT
TGATATGCTTATACACTTACACTTTATGCACAAAATGTAGGGTTATAATAATGTTA
ACATGGACATGATCTTCTTTATAATTCTACTTTGAGTGCTGTCTCCATGTTTGATG
TATCTGAGCAGGTTGCTCCACAGGTAGCTCTAGGAGGGCTGGCACCTTAGAGGTGG
GGAGCAGAGAATTCTCTTATCCAACATCAACATCTTGGTCAGATTTGAACTCTT

~~Sequence ID G6~~SEQ ID NO: 497

GGATTTTTGGTCCGCACGCTCCTGCTCCTGACTCACCGCTGTTTCGCTCTCGCCGAG
GAACAAGTCGGTCAGGAAGCCCGCGCGCAACAGCCATGGCTTTTAAGGATACCGGA
AAAACACCCGTGGAGCCGGAGGTGGCAATTCACCGAATTCGAATCACCTAACAAG
CCGCAACGTAAAATCCTTGGAAGGTGTGTGCTGACTTGATAAGAGGCGCAAAAG
AAAAGAATCTCAAAGTGAAAGGACCAGTTCGAATGCCTACCAAGACTTTGAGAATC
ACTACAAGAAAACTCCTTGTGGTGAAGGTTCTAAGACGTGGGATCGTTTCCAGAT
GAGAATTCACAAGCGACTCATTTGACTTGCACAGTCCTTCTGAGATTGTTAAGCAGA
TTACTTCCATCAGTATTGAGCCAGGAGTTGAGGTGGAAGTCACCATTGCAGATGCT
TAAGTCAACTATTTTAATAAATTGATGACCAGTTGTTAAAA

~~Sequence ID -- 61~~ SEQ ID NO: 498 nt: 362

CTTATTGAAAATTTTACTAATTTCTTACTTTTTAGGTTTTAGGAGAATACTTTTGGG
TAATTGACTAGCCTCACATTATATTGATAGAGGTTCTTGAAAACTTTAATGCCAAT
TCATGTATCTTATGACTAAAATAGATAATCCATTTAGAAATTTAAGTCATTCTTGC
GTGCTTGATATGTGTCAGCACTATCCAAGTTGCTAGGGGATACAATGGTGAAGTG
AAAATATCAGCTAGGTGCCGGTGGCTCACACCTGTTATCCCAACAGTTTGGGAGG
CCAGGGTGGGAGGATCACTCAAGCACANGCGTTTCACACCAGCCTGGACAACAT
ACAAGACCCCATCTTTACCAAAGTTAAG

~~Sequence ID -- 490~~ SEQ ID NO: 500 nt: 382

TTTTCTTAGAACTTTATTTTTTCTGGCCAGGCGCAGTGGCTCACACCTGTAATCCC
AGCACTTTGGGAGGCCAAGGCAGGTCGATCACCTGAGGTCAGGAGCTCAAGACC
AGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCTGG
GCGTGGTGGCGCATGCCTGTAATCCCANCTACTCAGGAGGCTGAGGCAGGAGAA
TTGTTTGAACCCGGGAGGCGGAGGTTGCANTGAGCCGAGATTGCGCCACTGCACT
CCAGCCTGGGCAACAGAGCGAACTCCATCTCAAAAAAAAAAAAAAAAAACAAC
CTTTATTTTTTCTGATTTTAAAAGTAATAACTAGTTTGTAGAAACATTAAAAGT

~~Sequence ID -- 892~~ SEQ ID NO: 501 nt: 559

TCTTTCGGAAGCGCGCCTTGTGTTGGTACCCGGGAATTCGCGGCCGCGTCGACGC
GGTCGTAAGGGCTGAGGATTTTTGGTCCGCACGCTCCTGCTCCTGACTCACCGCT
GTTTCGCTCTCGCCGAGGAACAAGTCGGTCAGGAAGCCCGCGCGCAACAGCCATG
GCTTTTAAGGATACCGGAAAAACACCCGTGGAGCCGGAGGTGGCAATTCACCGA
ATTCGAATCACCCCTAACAAGCCGCAACGTAAAATCCTTGAAAAGGTGTGTGCTG
ACTTGATAAGAGGCGCAAAAGAAAAGAATCTCAAAGTGAAAGGACCAGTTTCGAA
TGCCTACCAAGACTTTGAGAATCACTACAAGAAAACTCCTTGTGGTGAAGGTTT
TAAGACGTGGGATCGTTTCCAGATGAGAATTCACAAGCGACTCATTGACTTGAC
AGTCCTTCTGAGATTGTTAAGCAGATTACTTCCATCAGTATTGAGCCAGGAGTTG
AGGTGGAAGTCACCATTGCAGATGCTTAAGTCAACTATTTTAATAAATTGATGAC
CAGTTGTTT

~~Sequence ID -- 77~~ SEQ ID NO: 499 nt: 464

GCGGCTGCTGTTGGTTGGGGGCCGTCCCGCTCCTAAGGCAGGAAGATGGTGGCCG
CAAAGAAGACGAAAAAGTCGCTGGAGTCGATCAACTCTAGGCTCCAACCTCGTTAT
GAAAAGTGGGAAGTACGTCCTGGGGTACAAGCAGACTCTGAAGATGATCAGACA
AGGCAAAGCGAAATTGGTCATTCTCGCTAACAACCTGCCAGCTTTGAGGAAATCT

GAAATAGAGTACTATGCTATGTTGGCTAAAACTGGTGTCCATCACTACAGTGGCA
ATAATATTGAACTGGGCACAGCAGCATGCGGAAAATACTACAGAGTGTGCACACTGG
CTATCATTGATCCAGGTGACTCTGACATCATTAGAAGCATGCCAGAACAGACTGG
TGAAAAGTAQAACCTTTTCACCTACAAAATTTACCTGCAAACCTTAAACCTGCAA
AATTTTCCTTTAATAAAATTTGCTTG

Abstract

Product and Method

The present invention relates to oligonucleotide probes, for use in assessing gene transcript levels in a cell, which may be used in analytical techniques, particularly diagnostic techniques and kits containing the same.